OCT 27 1924

VOLUME 24 SECOND SERIES NUMBER 4

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

THE PHYSICAL REVIEW

A JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS

CONDUCTED BY

THE AMERICAN PHYSICAL SOCIETY

BOARD OF EDITORS

Managing Editor: Gaertner S. Fulcher, Corning, New York

E. P. Adams, A. J. Dempster, A. L. Hughes, A. H. Compton, E. H. Kennard, Leigh Page, H. D. Smyth, W. G. Cady, John T. Tate

OCTOBER, 1924

Ionizing Potentials of Multiatomic Gases – – – – – – – – – – – – C. A. Mackay 329
Absorption of Radiation by Multiply Periodic Orbits – – – – – – – J. H. Van Vleck 330
Part I. Some Extensions of the Correspondence Principle – – – – – – – – – – 330
Part II. Calculation of Absorption by Multiply Periodic Orbits – – – – – – – – 347
Variation of Thermionic Emission with Temperature – – – – – – – – A. T. Waterman 395
Change of Conductance of Selenium due to Electronic Bombardment Ralph De Lara Kronig 397
Diffusion of Alkali Salt Vapors in the Bunsen Flame – – – – – – – – Gardner E. Davis 399
Electrostatic Method for Measuring Dielectric Constants of Liquids – – – – – A. P. Carman 405
Electric Moment of Gaseous Molecules of Halogen Hydrides – – – – – – – C. T. Zahn 409
Magnetic Susceptibility of Helium, Neon, Argon, and Nitrogen – – – – – – L. G. Hector 418
Crystal Structure of Hard Steel – – – – – – – – – – – – – – K. Hessenbruch 425
Resilience Effect for Silicon Steel – – – – – – – – – – – – – E. P. T. Tyndall 430
Speed of Sound in Sea Water – – – – – – – – – – – – – – – – E. A. Eckhardt 435
Partials of a Pianoforte String Struck by an Elastic Hammer – – – – – – – R. N. Green 436
Book Reviews – 451

PUBLISHED BY THE AMERICAN PHYSICAL SOCIETY
CORNING, N. Y., AND MENASHA, WIS.

THE AMERICAN PHYSICAL SOCIETY

OFFICERS OF THE SOCIETY

President:	C. E. MENDENHALL	Univ. of Wisconsin, Madison, Wis.
Vice-Presidents	G. C. MILLER	Carnegie Institution of Applied Science, Cleveland, Ohio
Secretary	H. W. WEBB	Columbia University, New York City
Treasurer	G. B. PEGRAM	Columbia University, New York City
Editor	G. S. FULCHER	Cornell University, Ithaca, N. Y.
Local Secretary for the Pacific Coast	S. L. WEBSTER	Stanford University, Calif.

THE PHYSICAL REVIEW

Publication monthly. Two volumes annually.

First series. From July 1, 1893 to January 1, 1913, the Review was conducted by Edward L. Nichols, Ernest Merritt and Frederick Bedell, with the financial support of Cornell University. The first four volumes were issued bimonthly; subsequent volumes, monthly. Thirty-five volumes, in all.

Second series. On January 1, 1913, the Review was transferred to the American Physical Society as announced on the first page of this new series which was then begun. Since then it has been conducted by a Board of Editors elected by the Society. Two annual volumes begin in January and July.

Subscriptions are referred to page 2 of the advertising pages for further information.

Annual Subscription, throughout the Postal Union, $7.00 in advance.

Single Copies, $.75.

Foreign Duties, 1909-1918 inclusive, in advance.

Back Volumes, $6.50 per volume; Back Numbers, $1.25 per copy. Postage is prepaid on orders accompanied by remittance; otherwise postage is extra. The Second volumes are complete. The Review offers to purchase a limited number of First Series, Vol. 1, No. 1 (at $1.25 each); Second Series, Vol. 7, No. 1, January 1916; Vol. 8, No. 1 and Vol. 9, No. 3, February 1917 (at $1.50 each); Vol. 13, No. 1, January 1919 ($.75 each).

Advertisements. Terms will be supplied by the Editor on request. Copy for each number must be received at Cornell before the 15th of the preceding month; old copy may be used until new copy is received. Make all checks payable to the publisher, George Banta Publishing Co., Menasha, Wisconsin.

Discussion of any paper is allowed subscription agents and members of the American Physical Society.

Communications relating to missing numbers, to receipt of the Review, to change of address, etc., should be addressed to the Treasurer; otherwise, to the Editor, Corning, N. Y.

Correspondence should be addressed to the Secretary, Corning, N. Y.

Entered at the postoffice [illegible] as second-class matter [illegible] Act of August 24, 1912 [illegible] Act of June 30, 1906 [illegible] Act of October 3, 1917 [illegible] authorized September, 1919

Second Series *October, 1924* *Vol. 24, No. 4*

THE

PHYSICAL REVIEW

IONIZING POTENTIALS OF MULTIATOMIC GASES

By C. A. Mackay

Abstract

Ionizing potentials for helium and fifteen compound inorganic gases were measured by accelerating photo-electrons through a gauze into a chamber where the positive ions produced were drawn to a fine Pt electrode, made small so as to eliminate effects due to radiation. The current-voltage curves obtained consist of straight lines except at the breaks. Corrections for initial velocity, etc., were made by obtaining on the same curve both a break for the unknown gas and one for mercury (10.4 volts). The values found are, in volts: *helium*, 24.5; *hydrogen*, 15.8; *nitrogen*, 16.3; *oxygen*, 12.5, 16.1; *HCl*, 13.8; *HBr*, 13.2; *HI*, 12.8; *water*, 13.2; NH_3, 11.1; Cl_2, 13.2; Br_2, 12.8; I_2, 10.0; *NO*, 9.4; CO_2, 14.3; *CO*, 14.1, 15.6; H_2S, 10.4. Comparison is made with results of other observers. ***Theoretical interpretation.*** An attempt is made to test the thermochemical method of estimating ionizing potentials in those cases (H_2S, NH_3, H_2O) where data are available. No process of ionization involving molecular dissociation seems consistent with the experimental results. This and other evidence suggests that in those cases (HCl, HBr, HI) in which the thermochemical calculations yield accurate results, the agreement may be fortuitous, and that the lowest ionization potentials in each case corresponds to ionization without dissociation.

Introduction

In recent years the study of the critical potentials of monatomic gases and vapors has been of great importance. The experimental data available at present are extensive and accurate, and fruitful correlations between electrical and spectroscopic measurements have been found which have increased our knowledge of the structure of the atom and of the process of radiation and ionization. In the case of multiatomic molecules, not nearly so many measurements have been made and the observed values vary from one another over a much greater range than can be accounted for by experimental error. In much of the earlier work in this field, the distinction between radiating and ionizing potentials was not clearly drawn,[1] which, of course, added to the uncertainty of the results. The practically unknown structure of the molecule has made

[1] Davis and Goucher, Phys. Rev. **10**, 101 (1917)

the interpretation of the experimental results very difficult. The greatest progress has been along thermochemical lines,[2] especially in the case of the halogens and some of their salts. The application of the positive ray method to molecular ionization[3] has also proved valuable, since it gives definite experimental knowledge of the various products of ionization, which have to be assumed in any other method of measurement.

The present investigation was undertaken for two reasons; to measure the ionizing potentials of as large a number of multiatomic gases as possible, so that more data would be available; and to test the application of thermochemical reasoning to additional gases. It was thought that

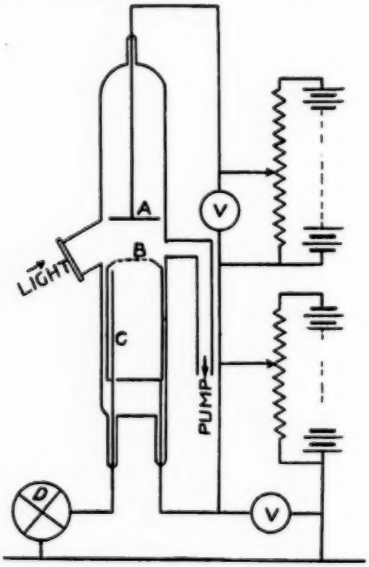

Fig. 1. Diagram of apparatus: *A*, nickel target; *B*, platinum gauze; *C*, fine wire electrode; *D*, electrometer; *VV*, voltmeters

interesting results might be found, especially in compounds which are non-polar, and that, in connection with other lines of work, our knowledge of the process of molecular ionization might be increased.

Method

The experimental tube and the electrical connections are shown in Fig. 1. Light from a quartz mercury arc fell on the nickel target *A*. The electrons emitted by the target were accelerated towards the gauze *B* by a variable difference of potential. Ionization took place in the region

[2] A summary is given in Foote and Mohler, Origin of Spectra, Chapter VIII.

[3] H. D. Smyth, Phys. Rev. **14**, 409 (1919); Proc. Roy. Soc. A **102**, 283 (1922); **104**, 121 (1923); **105**, 117 (1924)

beyond B and the positive ions were collected on the electrode C by means of a small constant difference of potential between B and C of either 4 or 6 volts. C was made of fine platinum wire, so that its area was reduced to such an extent that the effect of radiation on C was practically eliminated. For more complete details and for the method of correcting for the initial velocity of the electrons when they leave the surface of the target, the reader is referred to a previous paper.[4]

OXYGEN

The results have already been given for nine gases[4] and it has been shown that in every case except oxygen the agreement with the more recent determinations has been good. The work on oxygen has been repeated, using the same experimental tube, but preparing the oxygen in

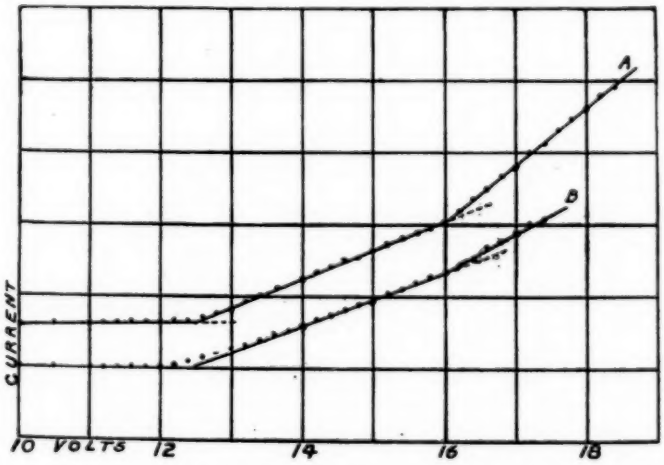

Fig. 2. Typical current-voltage curves for oxygen, with nickel target as a source of electrons (A); with a hot filament as a source of electrons (B).

two ways; first, by the electrolysis of sulphuric acid and, later, by heating potassium chlorate. Hydrogen was removed by red hot copper oxide and a phosphorous pentoxide tube was used to dry the gas. A second series of runs was made in which the target A was replaced by a small molybdenum filament, which was electrically heated to a dull red. The latter method eliminated the possible presence of ozone which might have been produced by the action of the ultraviolet light on the oxygen.

Fig. 2 shows typical current-voltage curves in which the necessary correction in the voltage has been made. In these curves, as in all the others, the current scale is purely arbitrary. The reciprocal of the rate of

[4] Mackay, Phil. Mag. **46**, 828 (1923)

deflection of the electrometer measured the current and it was plotted on a convenient scale. The lower break at 12.5 volts is present whether the source of electrons is the nickel target or the hot filament. In addition there is a clearly marked discontinuity at 16.0 volts. Both these values must be attributed to molecular oxygen, since the moderately heated filament could have produced no appreciable dissociation. 16.1 volts has been found by Lockrow[5] to be the lowest value at which an arc can be maintained.

Hydrogen Sulphide

The observed value for the ionizing potential was 10.4 volts. Fig. 3 shows a typical curve B, in which the accelerating voltage was varied

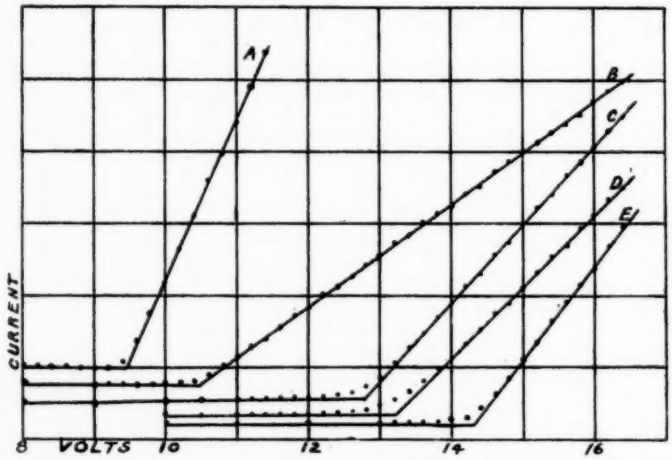

Fig. 3. Curves for nitric oxide (A); hydrogen sulphide (B); bromine (C); chlorine (D); carbon dioxide (E).

from 0 to 17 volts. The absence of any break at 15.8 volts shows that there was very little if any dissociation of the hydrogen sulphide by the ultraviolet light, so that the discontinuity at 10.4 volts is the result of ionizing the molecule of the compound. The single break is also an indication that no large proportion of the molecules is dissociated by electron impact. This point will be discussed in more detail in a later section.

Carbon Dioxide, Nitric Oxide and Carbon Monoxide

Carbon dioxide and nitric oxide gave curves very similar to that for hydrogen sulphide. They are shown in Fig. 3, curves E and A respectively. Carbon monoxide has two ionizing potentials, both of which are

[5] Lockrow, Abstract of paper presented at New York Meeting Am. Phys. Soc. February, 1924, Phys. Rev. **23**, 553 (1924)

clearly marked in the typical curve B in Fig. 4. There is no evidence of dissociation. It is interesting to note that carbon monoxide is one of the few molecules which occurs with two positive charges in positive ray analysis.[6] The present method gives no direct information regarding the products of ionization, but in this case it is at least probable that the two values of the ionizing potential correspond to the loss of one and two electrons from the molecule.

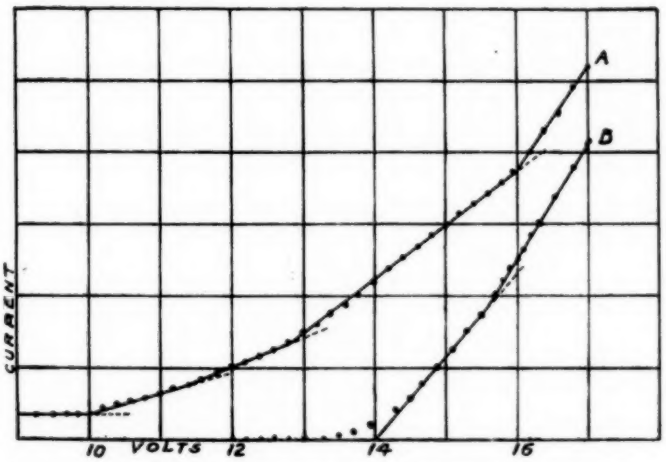

Fig. 4 Curves for hydrogen iodide (A) and for carbon monoxide (B)

Chlorine, Bromine and Iodine

The vapor pressures of chlorine and bromine are high enough to permit direct measurement of their ionizing potentials, while for iodine an indirect method was employed. The type of curve obtained for chlorine is shown in Fig. 3 curve D. Bromine is very similar, as appears from curve C.

To obtain the ionizing potential of iodine, hydrogen iodide vapor was placed in the experimental tube. Fig. 4 curve A shows the result between 0 and 17 volts. Four breaks are quite apparent. To show the higher breaks more clearly, after the accelerating potential had been increased from 0 to 14 volts, the intensity of illumination on the target (Fig. 1) was reduced by inserting copper gauze screens and the run was continued from 14 volts on, and curve B was obtained. It was observed that when readings were made immediately after the light was admitted to the tube, the first ionization occurred in the neighbourhood of 13.0 volts. Prolonged exposure of the gas gave the type of curve shown, with the first ionization at 10.0 volts. Guided by this fact and by the results of

[6] J. J. Thomson, Positive Rays, pp. 84 and 218.

previous measurements[7] on hydrogen iodide and iodine, the interpretation of the four breaks is clear. The lowest is due to the pressure of iodine vapor, produced by the dissociation of hydrogen iodine, while that at 12.8 volts is given by the undissociated molecule. The other two are very probably to be attributed to hydrogen, and they agree fairly well with the results of other investigations.[8]

For the convenience of the reader the values for all the gases thus far measured are given in Table I. Some of the more recent measurements by other observers have been added, but no attempt has been made to

TABLE I

Ionizing potentials of multiatomic gases

Gas	Observed values	Values observed by others	References
H_2	15.8	16.0	Olmstead[8]
		15.8	Davis and Goucher[1]
		16.0	Foote and Mohler, Origin of Spectra, p. 68.
N_2	16.3	16.2	Duffendack, Phys. Rev., **20,** 665 (1922).
		16.9	Foote and Mohler, J. Opt. Soc. Am. **4,** 49 (1920).
		17.0	Brandt, Zeit. f. Phys. **8,** 32 (1921).
O_2	12.5, 16.1	15.5	Foote and Mohler, Bur. Stds. Sci. Paper, No. 400 (1920)
		14.0	Boucher, Phys. Rev. **19,** 189 (1922).
		15.5	Smyth, Proc. Roy. Soc. **A105,** 116 (1923).
HCl	13.8	13.7	Knipping[7]
			Franck, Zeit. f. Phys. **11,** 160 (1920).
HBr	13.2	13.3	Franck, loc. cit.
HI	12.8	12.7	Franck, loc. cit.
H_2O	13.2	13.0	Foote and Mohler, Phys. Rev. **17,** 394 (1921).
NH_3	11.1		
He	24.5	24.5	Franck and Knipping, Zeit. f. Phys. **1,** 320 (1920) Also Franck, loc. cit.
		24.5	Horton and Davies, Roy. Soc. Proc. **97,** 23 (1920).
			Also Davies, Phil. Mag. **45,** 797 (1923).
Cl_2	13.2		
Br_2	12.8	11.3	Mohler (unpublished)
I_2	10.0	9.4	Smyth and Compton[7]
		10.5	Foote and Mohler, Bur. Stds. Sci. Paper No. 400, (1920).
NO	9.4	9.3	Hughes and Dixon, Phys. Rev. **10,** 495 (1917).
CO_2	14.3		
CO	14.1, 15.6	14.3	Foote and Mohler, Phys. Rev. **17,** 394 (1921).
		14.5	Stead and Gessling, Phil. Mag. **40,** 413 (1920).
		15.1	Found, Phys. Rev. **16,** 41 (1920).
H_2S	10.4	(9.3?)	Hughes and Dixon, Phys. Rev. **10,** 495 (1917).

[7] Foote and Mohler, Jour. Amer. Chem. Soc. **43,** 1832 (1920);
Knipping, Zeits. f. Phys. **7,** 328 (1921);
Smyth and Compton, Phys. Rev. **16,** 501 (1920).

[8] Olmstead, Phys. Rev. **20,** 613 (1922)

give a complete summary of the data. A complete summary of experimental results and their interpretation is soon to appear in a Bulletin of the National Research Council, written by K. T. Compton and F. L. Mohler.

Discussion of the Results

The quantum theory has made possible very many correlations between spectroscopic lines and critical potentials for the *atoms* of many elements. When the band spectrum of the multiatomic *molecule* is considered, the problem becomes greatly complicated, since we may have energy transfers in the electronic, vibratory, or rotational quantized states of the molecule. While the critical potential method[9] of attack has been employed to interpret radiating potentials, for ionizing potentials a consideration of the thermochemical relations has led to more direct results, at least in the case of certain polar compounds. Calculations by means of this method are given by Foote and Mohler[10] for the three halogen acids and the theoretical values agree well with the experimental data.[11] For other types of compounds the agreement is not so satisfactory. Born and Bormann[12] have predicted 31.8 volts for the complete ionization of hydrogen sulphide, the final products being assumed to be $2(H)^+$ and S^{--}. The equation they used may be written, according to Foote and Mohler,[13] in the following way:

$$[H_2S]+Q(H_2S)+2D(H_2)+2J(H)+S(S)-E(S)\rightarrow 2[H]^+ + S^{--}$$

where square brackets denote the gaseous phase, $Q(H_2S)$ is the heat of formation of H_2S, $D(H_2)$ is the heat of dissociation of H_2 per 1/2 gram molecule, J is the ionizing potential, S is the heat of sublimation per gram molecule, E is the electron affinity of one gram atom. From this equation we get at once

$$J(H_2S)=Q(H_2S)+2D(H_2)+2J(H)+S(S)-E(S)$$

which gives $J(H_2S)=31.8$ volts if we use Born and Bormann's values for quantities on the right side of the equation. This is far above the lowest observed ionizing potential so that it is probable that the first ionization does not involve the loss of both electrons.

[9] Foote and Mohler, Origin of Spectra, p. 190.
[10] Foote and Mohler, Origin of Spectra, p. 186.
[11] Mackay, loc. cit.[4]
[12] Born and Bormann, Zeits. f. Phys. **1**, 250 (1920)
[13] Foote and Mohler, Origin of Spectra, pp. 180-186.

There are several possibilities, for which sufficient data are available to get at least approximate results. The process may be simply

$$[H_2S]+J(H_2S)\rightarrow[H_2S]^++e$$

where e denotes a free electron, and $[H_2S]^+$ is a molecule of hydrogen sulphide with a single positive charge. No thermochemical quantities are involved here since we have no dissociation. Or, we may assume that we have dissociation but not double ionization and write an equation thus

$$[H_2S]+Q(H_2S)+S(S)+2D(H_2)+J(H)-E(S)^--Q(H+S^-)\rightarrow H^++[HS]^-,$$

which gives

$$J(H_2S)=Q(H_2S)+S(S)+2D(H_2)+J(H)-E(S)^--Q(H+S^-)$$

In this equation the quantities are the same as in Born and Bormann's with the exception of the last two. $E(S)^-$ means the electron affinity of sulphur for one electron and $Q(H+S^-)$ is the heat of formation of the hypothetical compound $(HS)^-$ from neutral hydrogen atoms and singly charged sulphur atoms. As to $E(S)^-$, it is probable that it has a value between $E(S)^{--}$ and 0 and nearer the former than the latter; that is, nearer 2.0 volts than 0. For an approximate value, let us assume $E(S)^-$ is 1.5 volts. $Q(H+S^-)$ is also unknown. On Bohr's theory, S^- would have the same number of electrons in the outer shell as the atom of chlorine, but probably the tendency to complete the shell would be greater in the case of chlorine than for S^-. This gives us a means to approximate the value of $Q(H+S^-)$ by making it equal to or less than $Q(HCl)$. The value given in the tables of Landolt-Börnstein for the latter is 5.24 volts, so that $Q(H+S^-)$ is probably less than 5.2 volts. Given the two quantities $E(S)^-$ and $Q(H+S^-)$ we can substitute all the values in the equation given above, and obtain

$$J(H_2S)=.2+2.9+3.6+13.5-1.5-5.2=13.5 \text{ volts.}$$

Even if the approximations made were considerably in error, we have a difference of 3.1 volts between the calculated and the observed values of the ionizing potential when we use the most favorable estimate of $E(S)^-$ and $Q(H+S^-)$. Probably the discrepancy between calculation and experiment should be greater. In exactly the same way it may be shown that any of the following products of ionization, H_2+S^+ or H^++H+S, or $H_2^++S^-$ lead to an ionizing potential still higher than 13.5 volts. From these considerations one concludes that the lowest observed ionization (10.4) consists in the removal of an electron from the hydrogen sulphide molecule and that *dissociation does not take place.* As we have already

seen, this is supported by the experimental curve, which shows only a single ionization up to 17 volts.

Both ammonia and water may be treated in the same way as hydrogen sulphide. For the former we may write $NH_3+J(NH_3)\rightarrow NH_3^+ + e$ as the simplest possible equation. If, on the other hand, we assume that N, H_2 and H^+ are formed we have

$$NH_3+Q(NH_3)+D(N)+D(H)+J(H)\rightarrow N+H_2+H^+ + e\ .$$

We do not know $D(N)$, but Langmuir's[14] work shows that it is greater than $D(H)$; that is, greater than 1.8 volts. This gives us the equation:

$$J(NH_3) > .4+1.8+1.8+13.5 = 17.5 \text{ volts},$$

where $Q(NH_3)$ is taken from Landolt-Börnstein tables. In the same way it may be shown that any assumed process of ionization which does not involve the electron affinity of any atom or molecule will give an ionizing potential several volts above the experimental result. From the chemical properties of hydrogen and nitrogen, it seems unlikely that there is any molecule or atom derived from ammonia for which the electron affinity is more than a fraction of a volt. We are left with the same conclusion as in the case of hydrogen sulphide, that the first ionization is due to the expulsion of an electron and that the positive ion is $(NH_3)^+$.

Turning our attention to water vapor, we find that the data here are not sufficiently well known to do much with thermochemical equations. The oxygen atom is known to have an electron affinity but its magnitude is still undetermined. The same holds true for the OH group which may be one of the ions resulting from ionization. We are limited by these unknown quantities to consideration of those cases in which the electron is supposed to be liberated. A possible equation may be written thus

$$H_2O+Q(H_2O)+J(H_2)\rightarrow H_2^+ + O + e\ ,$$

and from this we get

$$J(H_2O) = Q(H_2O)+J(H_2) = 2.3+15.8 = 18.1 \text{ volts}\ .$$

If the oxygen atom has an electron affinity of 5.0 volts, then the calculated result would be reduced by that amount and would be in agreement with the observed value, 13.0 volts. Another possible equation is

$$H_2O+Q(H_2O)+2D(H_2)+J(H)\rightarrow H^+ + H + O + e\ ,$$

whence

$$J(H_2O) = 2.3+3.6+13.5 = 19.4 \text{ volts}\ .$$

[14] Langmuir, Jour. Amer. Chem. Soc. **34**, 860 (1912); **37**, 417 (1915)

Neither of these equations seems to hold since both give too high an ionizing potential and there is no evidence of any ionization of free oxygen or hydrogen showing in the curve for water vapor. The more probable supposition is that water vapor behaves in the same manner as hydrogen sulphide.

In addition to the negative evidence obtained from considerations similar to those just outlined, there is positive evidence as to the nature of the first ionization. Smyth[3] has studied hydrogen, oxygen, and nitrogen by means of positive ray analysis. He has found that for all three gases the lowest ionization is not accompanied by dissociation, unless as a secondary process occurring when the ionized or excited molecule later collides with a neutral molecule and either itself dissociates or effects the dissociation of the impacted molecule. He states that in view of his results, the validity of thermochemical reasoning, when applied to non-polar compounds, is extremely doubtful.

In the case of polar compounds like the halogen acids and salt vapors, the case is also doubtful. On the one hand we have the satisfactory agreement of experiment with thermochemical calculations in the case of the halogen acids[15] and also the recent detection of I^- and Cl^- in the products of ionization of HgI_2 and $HgCl_2$ in the work of Kondratjeff and Semenoff,[16] who used Smyth's positive ray method of analysis of products of ionization. These authors thus detected I^- and Cl^- at voltages which agreed, within the limits of experimental error, with the values predicted from thermochemical considerations. The possible error, however, was several volts, and, furthermore, it remains to be discovered whether these negative ions were primary products of ionization or secondary products arising from subsequent collisions with neutral molecules. This point is now being investigated by Smyth and Barton for several halogen compounds.

In this connection it may be noted that Duffendack,[17] in intense arcs in HCl vapor up to 70 volts, has failed to observe the hydrogen spectrum. This is very surprising if the products of ionization are H^+ and Cl^-, as assumed in making thermochemical calculations.

In conclusion it may fairly be stated that the experimental evidence is strongly against dissociation with ionization as a primary process at the lowest ionizing potential of non-polar compounds and that, in the case of polar compounds, there is still a question as to whether the close agree-

[15] Foote and Mohler, loc. cit.[10]

[16] Kondratjeff and Semenoff, Zeit. f. Phys. **22**, 1 (1924)

[17] Duffendack, unpublished

ment between experiment and thermochemical calculation is not fortuitous, and whether such dissociation as has been observed may not be a secondary process.

The writer takes this opportunity to thank Professor K. T. Compton for his most valuable suggestions and unfailing interest throughout the course of this work.

PALMER PHYSICAL LABORATORY,
PRINCETON UNIVERSITY,
June 17, 1924.

THE ABSORPTION OF RADIATION BY MULTIPLY PERIODIC ORBITS, AND ITS RELATION TO THE CORRESPONDENCE PRINCIPLE AND THE RAYLEIGH-JEANS LAW.

PART I. SOME EXTENSIONS OF THE CORRESPONDENCE PRINCIPLE

By J. H. Van Vleck

Abstract

This part deals with the quantum theory aspects of the problem. In the absence of external radiation fields the distortion in the shape of the orbit is essentially the same in both the classical and quantum theories provided in the former we retain only one particular term τ_1, τ_2, τ_3 in the multiple Fourier expansion of the force $2e^2\ddot{v}/3c^3$ on the electron due to its own radiation. The term to be retained is, of course, the combination overtone asymptotically connected to the particular quantum transition under consideration. Then the changes $\Delta J_1, \Delta J_2, \Delta J_3$ in the momenta J_k which fix the orbits and which in the stationary states satisfy the relations $J_k = n_k h$, are in the ratios of the integers τ_1, τ_2, τ_3 in both the classical and quantum theories, making the character of the distortion the same in both even though the speed of the alterations may differ. One particular term in the classical radiation force is thus competent to bring an orbit from one stationary state to another.

The correspondence principle is then extended so as to include absorption as well as the spontaneous emission ordinarily considered. Commencing always with a given orbit it is possible to pair together the upward and downward transitions in such a way that in each pair the upward and downward optical frequencies (determined by the $h\nu$ relation) are nearly equal for large quantum numbers (usually long wave-lengths). That is, if s denotes the initial orbit there exist levels r and t such that the ratio $(W_r - W_s)/(W_s - W_t)$ or ν_{rs}/ν_{st} approaches unity when the quantum numbers become large. We shall define as the differential absorption the excess of positive absorption due to the upward transition $s \rightarrow r$, over the negative absorption (induced emission) for the corresponding downward transition $s \rightarrow t$. It is proved that for large quantum numbers the classical theory value for the ratio of absorption to emission approaches asymptotically the quantum theory expression for the ratio of the differential absorption to the spontaneous emission. Consequently a correspondence principle which makes the numerical values of the emission in the two theories agree asymptotically, of necessity achieves a similar connection for the absorption.

The correspondence principle basis for a dispersion formula proposed by Kramers, which assumes the dispersion to be due not to the actual orbits but to Slater's "virtual" or "ghost" oscillators having the spectroscopic rather than orbital frequencies, is then presented. Kramers' formula has both positive and negative terms and the differential dispersion may be defined in a manner analogous to the differential absorption. It is shown that the quantum differential dispersion approaches asymptotically the dispersion which on the

classical theory would come from the actual multiply periodic orbit found in the stationary states. This asymptotic connection for the general non-degenerate multiply periodic orbit must be regarded as an important argument for Kramers' formula.

1. INTRODUCTION

ACCORDING to the first postulate of the Bohr theory of atomic structure the electrons can move only in certain particular quantized non-radiating orbits or stationary states. In order that the quantum conditions may be applicable it is necessary for the motion to be of the so-called "multiply periodic" type, which can be represented by multiple Fourier series of the form

$$x = \Sigma_{\tau_1\tau_2\tau_3} X(\tau_1,\tau_2,\tau_3) \cos\left[2\pi(\tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3)t + \gamma_{\tau_1\tau_2\tau_3}{}^{(x)}\right] \qquad (1)$$

with similar expansions by y and z. The constants ω_1, ω_2, ω_3 are the intrinsic orbital frequencies, and the summation is to be extended over all possible positive and negative integral values of the integers τ_1, τ_2, τ_3 subject to the restriction that $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$ be positive. For simplicity in notation we have assumed that there are only three ω's, but in general the number of such frequencies can be equal to or less than the number of degrees of freedom of the atomic system to which the electron under consideration belongs. (The slight modifications necessary to extend the results of the present paper to systems with more than three frequencies will be discussed in section 17). We shall also suppose that the system is "non-degenerate" so that the number of degrees of freedom is equal to rather than greater than the number of frequencies. This means that in the case of three ω's, the problem may be thought of as one of a single electron moving in an asymmetrical three-dimensional static force field. It will also be assumed that the orbits conform to the classical mechanics and that all solutions of the differential equations of motion can be represented by series of the form (1), thus making the complete dynamical system multiply periodic, rather than merely certain particular families of orbits.

The quantum conditions for determining the size of the stationary states consist in equating to integral multiples of h a set of orbital constants J_1, J_2, J_3 defined by the relations

$$\omega_k = \partial W/\partial J_k (k = 1,2,3)\ ; \quad 2\overline{T} = J_1\omega_1 + J_2\omega_2 + J_3\omega_3; \qquad (2)$$

where W is the total energy and $\overline{T}$ is the average kinetic energy. We can

therefore set $J_k = n_k h$ $(k = 1, 2, 3)$, where n_1, n_2, n_3 are the integers or quantum numbers characteristic of a stationary state.[1]

According to the second postulate of the Bohr theory an electron may pass from one allowed orbit to another. The frequency ν_{rs} of the quantum of light thus radiated or absorbed is determined by the familiar relation

$$h\nu_{rs} = W_r - W_s , \tag{3}$$

where W_τ denotes the energy of the stationary state τ (with $\tau = r, s$, etc.). In the case of emission, r is the initial state and s the final state, while for absorption the significance of these symbols is just reversed, as r represents the higher energy level.

Considerable information concerning the probabilities of the various transitions between different orbits is furnished by Einstein's derivation[2] of the Planck radiation formula, $\rho(\nu) = 8\pi\nu^3 hc^{-3}/(e^{h\nu/kT} - 1)$, for the specific energy density of black body radiation. Einstein assumed:

(I) The number of atoms N_τ in the state τ is given by the statistical mechanics formula

$$N_\tau = Ne^{-W(\tau)/kT} / \Sigma_\tau e^{-W(\tau)/kT} \tag{4}$$

where N is the total number of atoms, and the sum is to be extended over all possible states.[3]

(II) The amount of energy emitted in a time Δt by transitions from the state r to the state s is represented by a formula of the type

$$\Delta E_{r\to s} = h\nu_{rs} N_r [A_{r\to s} + B_{r\to s}\rho(\nu_{rs})]\Delta t . \tag{5}$$

(III) The amount of energy absorbed by transitions from the state s to the state r is given by

$$\Delta E_{s\to r} = h\nu_{rs} N_s B_{s\to r}\rho(\nu_{rs})\Delta t . \tag{6}$$

(IV) There is to be statistical equilibrium (i.e. $\Delta E_{r\to s} = \Delta E_{s\to r}$) when the energy density has the characteristic black body distribution given by the Planck formula. Using Eqs. (3) and (4), one can verify that this equilibrium condition will be fulfilled provided the probability coefficients

[1] The quantum conditions have been stated above in what may be termed the correspondence principle form. Some readers may be more familiar with the formulation used by Sommerfeld and others, which consists in equating certain phase or "quantum" integrals to integral multiples of h, but the two methods can readily be shown to yield the same results in non-degenerate systems. See, for instance, appendix 7 of Sommerfeld's "Atombau."

[2] Einstein, Phys. Zeit. **18**, 121 (1917)

[3] The "a priori probabilities" p_τ which ordinarily appear in front of the exponentials in Eq. (4) have been omitted, as we are concerned with non-degenerate systems where all states have the same a priori probabilities, making the p_τ's in numerator and denominator cancel.

$A_{\tau\to s}$, $B_{\tau\to s}$, and $B_{s\to\tau}$, which are independent of the temperature and energy density, satisfy the relations

$$B_{s\to\tau} = B_{\tau\to s} = (c^3/8\pi h\nu_{\tau s}^3)A_{\tau\to s}\,. \tag{7}$$

Einstein's criterion of statistical equilibrium under black body radiation thus tells us how emission and absorption vary with the density, Eqs. (5) and (6), and how they are related to each other, Eq. (7), but still leaves undetermined the magnitude of the coefficient $A_{\tau\to s}$ in (7). To evaluate approximately the latter it is customary to resort to the correspondence principle, as explained below.

It is well known that according to classical electrodynamics an accelerated electron radiates energy continuously and that if in particular the orbit is of the form (1) then the light thus emitted should be resolved by spectroscopes into frequencies which are combination overtones $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$ of the orbital frequencies ω_1, ω_2, ω_3. According to the quantum relation (3) there is, however, no such immediate connection between the spectroscopic frequency and actual frequencies of motion, but it can easily be proved that if the quantum numbers n_1, n_2, n_3 of the states r and s differ from each other by τ_1, τ_2, τ_3 units respectively, so that $\Delta J_k=\tau_k h$, then in the region of high quantum numbers (usually also long wave-lengths) the optical frequency ν approaches asymptotically the combination overtone $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$. Since the classical and quantum mechanisms thus then give nearly the same numerical values for optical frequencies even though irreconcilably different in character, it is natural to assume that they give in the case of high quantum numbers the same numerical results for the relative intensities of different spectral lines, at least when there is no radiation field (i.e. $\rho(\nu)=0$). It is, however, to be clearly understood that the asymptotic connection of *frequencies* is a necessary mathematical consequence of the quantum conditions, and is hence, following Ehrenfest, best termed the correspondence *theorem* for frequencies.[4] On the other hand the existence of an analogous relation for the *intensities* of lines radiated when $\rho(\nu)=0$ must be regarded as an additional hypothesis, which we shall call the correspondence principle for emission. The latter is generally accepted not only because of its inherent reasonableness but also because of its excellent experimental verification in the "selection principle" and in the researches of Kramers and others on the intensities of Stark effect components.

[4] To prove the correspondence theorem for frequencies we need simply note that $h\nu_{\tau s}=W_\tau-W_s=\Delta W$. Nearly consecutive orbits of large quantum numbers differ but little from each other in relative size and we can then without great error replace the increment ΔW of the energy by the differential $dW=\Sigma^3_1(\delta W/\delta J_k)\Delta J_k$. Using Eq. (2) and the relation $\Delta J_k=\tau_k h$, we thus get the result that in the limit $\nu/(\tau_1\omega_1+\tau_2\omega_2+\omega_3)=1$.

To formulate analytically the correspondence principle for emission we need simply note that from the classical expression $(2e^2/3c^3)\dot{v}^2$ for the rate of radiation from an electron having a vector acceleration , it follows that the amount of energy radiated as light of frequency $\tau_1\omega_1 + \tau_2\omega_2+\tau_3\omega_3$ in the time interval Δt is[5]

$$\Delta E = (16\pi^4e^2/3c^3)\,(\tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3)^4\,[D(\tau_1\,,\,\tau_2\,,\,\tau_3)]^2\,\Delta t \qquad (8)$$

where $D^2 = X^2 + Y^2 + Z^2$. If we multiply by N_r to take into account the radiation from N_r electrons, then on comparing this classical expression with the quantum Eq. (5) for $\rho(\nu)=0$, and using the approximation $\nu=\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$, we see that the correspondence principle for emission requires that for high quantum numbers the probability coefficient $A_{r\to s}$ must have a value close to

$$A_{r\to s} = (16\pi^4e^2\nu^3_{rs}/3hc^3)\,[D^r(\tau_1\,,\,\tau_2\,,\,\tau_3)]^2\,. \qquad (9)$$

where D^r denotes D evaluated for an orbit of the same size and shape as that found in the stationary state r.

An equation of the precise form (9) can be expected to hold only asymptotically. At ordinary wave-lengths it is more probable that $A_{r\to s}$ depends not on the orbital frequencies and amplitudes evaluated for one particular stationary state but instead involves these quantities averaged in some manner or other over the continuous succession of "dis-allowed" orbits intermediate between the initial and final states. We might still have (9) valid in the limit if we assume, for instance, that a more exact expression is

$$A_{r\to s}=(16\pi^4e^2\nu^3/3hc^3)\left(\nu^{-m}\int_0^1(\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3)^m[D(\tau_1,\,\tau_2,\,\tau_3)]^{\pm2}d\lambda\right)^{\pm1}. \qquad (10)$$

Here m is some integer, while λ is an auxiliary parameter such that $J_k=(n_k+\tau_k\lambda)h$, where n_k and $n_k+\tau_k$ $(k=1,\,2,\,3)$ denote respectively the quantum numbers of the states s and r. Either the $+$ or $-$ sign must be consistently used throughout. Ordinarily the $+$ sign is taken, and then the special cases of $m=0,4$ have been studied in some detail by F. C. Hoyt.[6]

2. A Correspondence Principle for Orbital Distortions

In the review of the correspondence principle for emission given in section 1 it must be remembered that actually a classical electron radiates simultaneously all the combination overtones in the multiple Fourier

[5] For simplicity in printing, the arguments τ_1, τ_2, τ_3 of the amplitudes X, Y, Z, D will often be omitted; also the subscripts are often dropped from ν_{rs}.

[6] F. C. Hoyt, Phil. Mag. **46**, 135 (1923); **47**, 826 (1924). Eq. (10), of course, gives by no means all the possible expressions for $A(_{r\to s})$. In fact from certain viewpoints the best arguments appear to be for a certain type of logarithmic average first introduced by Kramers (The Intensities of Spectral Lines, Dan. Acad. Memoires, 1919, p. 330) and also studied by Hoyt. Only two out of six formulas studied by Hoyt are included in Eq. (10).

expansion rather than just one harmonic vibration component $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$. One can, however, formally avoid this difficulty by introducing what we shall term an "abridged" radiation force obtained by retaining only the one term τ_1, τ_2, τ_3, in the multiple Fourier expansion of the complete radiation force $2e^2\ddot{v}/3c^3$. It can then be verified that the total radiation is on the average given by Eq. 8.[7] We can then imagine the emitted light in the classical theory to be monochromatic, and its frequency will approach asymptotically the optical frequency ν in the quantum theory, while an asymptotic connection of the average rates of spontaneous radiation can be secured by using some such equation as (10).

A question which naturally arises at this point is whether the abridged radiation force with only one term produces just the same distortion in the size and shape of the orbit in the classical theory as the electron experiences in the quantum theory when it passes from one stationary state to another. In the classical theory an orbit having originally the same size and shape as the initial quantized stationary state will obviously after a properly chosen lapse of time be sufficiently damped by the abridged radiation force to make the energy the same as that of the final stationary state in the quantum theory. However, we cannot immediately infer that after this interval of time has expired the classical orbit will have the same shape as the final quantized orbit unless the system has only one degree of freedom, for then only does the energy determine uniquely the shape of the orbit. In the case, for instance, of an elliptical trajectory modified by a relativity precession to make the system nondegenerate (in two dimensions) the classical and quantum orbits might have initially the same semi-major axis and eccentricity; but when the classical orbit has been sufficiently damped by the abridged radiation force to make its energy (and hence approximately the semi-major axis) the same as that of another smaller stationary state, we cannot predict off-hand that then its eccentricity will become identical with that of the latter.

The writer is not aware of any specific statement in the literature as to whether these orbital distortions are the same in the two theories, except for a brief allusion to this question in an interesting article by H. A. Senftleben, which has just appeared.[8] However, an examination

[7] The average value of the work $-Fv\Delta t$ done in the time Δt against the abridged radiation force F has the value (8) even though the complete expansion of v (but not of $\ddot{v}$ in F) is retained, for "cross-terms" involving products of different combination overtones cancel out on the time average. For greater detail see section 16 of part II.

[8] H. A. Senftleben, Zeit. f. Phys. **22**, 127 (1924). Following Eq. (50) of his article Senftleben states the necessity of having the average time rate of change of the J's approach each other asymptotically in the two theories, and this implies an asymptotic

of this point is not difficult. In a system with 3 degrees of freedom the size and shape of the orbit are completely determined if we know the values of the expressions J_1, J_2, J_3 defined in (2), for the other arbitrary constants enter only as epoch angles. Now in the quantum theory the J's always change by integral multiples of h, and their alterations in going from one state to another are in the ratio

$$\Delta J_1 : \Delta J_2 : \Delta J_3 = \tau_1 : \tau_2 : \tau_3 . \tag{11}$$

On the other hand it is proved in section 16 of part II that if in the classical radiation force we keep only the term involving the combination overtone $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$, then the J's are also, in the classical theory, altered in the ratio given above for all time intervals. Therefore by including only the abridged force in the classical theory, the distortion is the same in the two theories. Another way of saying this is that one particular term in the classical radiation force damps the orbit in such a way that if at the start it coincides with a given quantized stationary state it tends to pass through a succession of stationary states of smaller energy, whose quantum numbers differ from those of the initial state by $m\tau_1$, $m\tau_2$, $m\tau_3$ units respectively, where m is an integer.

The results quoted in Eq. (11) hold even at ordinary wave-lengths (small quantum numbers) for it is shown in section 16 that according to the classical theory the J's always tend to change instantaneously in the ratio (11). Consequently the ΔJ's are in the ratio of integers even though the abridged radiation force be acting for such a long time that the alterations in the J's are of the order of magnitude of their initial values, making the radiated energy comparable with the total energy. The relations are also valid for systems with more than three J's (see section 17).

If we form a three-dimensional space for plotting values of J_1, J_2, J_3, the deformations of the orbit produced by one particular term in the classical radiation force may be represented by a straight line (neglecting small periodic fluctuations which cancel out on the average). The equations of these lines may be written $J_k = J_k^\circ + \lambda\tau_k h$, where λ is a parameter.

---

connection of the orbital distortions. No detailed proof of this, however, is contained in part I of Senftleben's article (the only part available at time of writing) but an indication is given of how it might be obtained from an asymptotic connection of the rates of radiation of angular momentum and energy in the two theories provided the system has only one electron. Both the method of proof and interpretation of the result (reached independently) in the present paper differ in many respects from the above, especially the ability to generalize the results to atoms with more than one electron (section 17), and the emphasis on the validity of the relation $\Delta J_1:\Delta J_2:\Delta J = \tau_1:\tau_2:\tau_3$ at ordinary wave-lengths.

Consequently the present considerations give a semi-theoretical basis for determining the probability coefficient $A_{r\to s}$ by averaging some function of the amplitudes and orbital combination frequencies along a straight line in the three dimensional space connecting the points corresponding to the initial and final states. This procedure, illustrated in Eq. (10), was adopted by Kramers and by Hoyt, but it is not apparent from the articles of these investigators whether they realized they were actually averaging along a path which would be traversed by the orbit under the influence of the abridged radiation forces. The various methods which have been suggested for determining $A_{r\to s}$ by averaging over the straight line path mentioned above are such as to make the mean free time in the initial state different from the actual time of transit between the two states under the influence of the abridged radiation force, except, of course, for an asymptotic connection of the two times for high quantum numbers. Neglecting higher powers of $1/c^2$, the quantum mean free time and the classical time of transit can be shown to be identical at ordinary wave-lengths or small quantum numbers provided in Eq. (10) we take the $-$ sign in both exponents and take the integer $m=-3$.[9] This result seems worth noting, but it is questionable how much significance should be attached to this method of procedure for there is no obvious reason why the two times should be exactly equal rather than asymptotically connected.[10]

3. Insufficiency of a Correspondence Principle Relating only to Emission

The correspondence principle for emission reviewed in section 1, cannot in a certain sense be regarded as entirely adequate because it establishes (or rather postulates) an asymptotic numerical connection of the classical

[9] To get the result quoted we need simply integrate Eq. (8) between $\lambda=0$ and $\lambda=1$, noting that $dE=(\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3)hd\lambda$ by Eq. (2). This gives the classical transit time. The quantum mean free time is on the other hand simply the reciprocal of $A_{r\to s}$.

[10] It is interesting to note that the classical time of transit from a 3_2 to a 2_2 orbit in hydrogen works out as 1.04×10^{-8} sec. Wien's canal ray experiments indicate a mean free time for Hα of 1.85×10^{-8} sec. (Ann. der Phys. **73**, 485, 1924). The latter, Wien shows, is almost exactly the reciprocal (1.87×10^{-8} sec.) of the logarithmic decrement for a linear oscillator of frequency ν (not an actual orbit), but this very close agreement is probably only a coincidence. On the other hand the value 1.04×10^{-8} sec. must be corrected to allow for the transitions $3_2\to2_1$, $3_1\to2_2$, which also contribute to Hα and have longer mean free times than $3_2\to2_2$,—the 3_1 state is, in fact, almost metastable—thus making the effective mean free time much larger. The agreement is as good as can be expected, for the correct formula for $A_{r\to s}$ probably does not make the quantum mean free time exactly equal to the classical time of transit, but instead involves a different kind of average of amplitudes and frequencies than that obtained by putting $m=-3$ in Eq. (10).

and quantum theory for intensity only in the special case that there is no radiation field (i.e. $\rho(\nu)=0$). The term present in Eq. (5) when $\rho(\nu)=0$ we shall call the spontaneous emission (Einstein's "Ausstrahlung"), while the second or remaining term proportional to the energy density we shall call the induced emission, although it is sometimes called the "negative absorption" in distinction from the true or positive absorption given by Eq. (6). The correspondence principle for emission correlates quantum theory spontaneous emission with classical theory emission due to the radiation force $2e^2\ddot{v}/3c^3$. On the other hand an electron can according to the classical theory absorb energy from a radiation field, and there must be some kind of an asymptotic connection between this absorption in the classical electrodynamics and the induced emission and absorption in the quantum theory. This will be discussed in sections 4 and 5. There can be no question of a correlation of the classical and positive quantum absorptions alone, as this would leave the induced emission unexplained. The existence of the induced emission term in the quantum theory may at first sight appear strange, but it is well known that this is qualitatively explained in that with the proper phase relations a classical electric wave may receive energy from an atomic system although on the average (i.e. integrating over all possible phase relations) it contributes more than it receives in exchange. It is therefore the excess of positive absorption over the induced emission which one must expect to find asymptotically connected to the net absorption in the classical theory.

4. Correspondence Principle for Absorption for a Linear Oscillator

Before seeking to develop a correspondence principle for absorption for the general case of an arbitrary multiply periodic orbit, we shall first for simplicity and clarity confine our attention to a one-dimensional linear oscillator. Here the multiple Fourier expansion reduces to

$$x = D\cos(2\pi\omega t + \gamma), \quad (y = z = 0).$$

As there is only one degree of freedom, there is just one quantum number n, and this can by the correspondence principle only change by one unit, as there are no harmonics of the fundamental frequency ω in the Fourier expansion. Now the energy W of a linear oscillator of amplitude D, mass m, and frequency ω is in general $2\pi^2\omega^2 mD^2$. Furthermore it is well known that for a linear oscillator the quantum conditions require that the energy W_n of a state of quantum number n have the value $nh\omega$. Therefore the amplitude D_n of an orbit of quantum number n is given by

$$D_n^2 = nh/(2\pi^2\nu m). \tag{12}$$

In writing (12) we have utilized the familiar fact that for a linear oscillator $\nu=\omega$, as here the spectroscopic and orbital frequencies are identical.[11] We can now apply the correspondence principle for emission, as embodied in Eq. (9), to determine the approximate value of the probability coefficient $A_{\tau\to s}$. Using the fact that in the present notation $D^n(\tau_1,\tau_2,\tau_3)$ is nothing but D_n, while ν_{rs} is simply ν, and taking r as an n quantum state, s as an $(n-1)$ quantum state, we thus get $A_{\tau\to s}=8\pi^2e^2\nu^2n/(3c^3m)$. From (7) it then follows that

$$B_{n\to(n-1)} = B_{(n-1)\to n} = n\pi e^2 / (3hm\nu). \tag{12a}$$

Hence by Eq. (6) we see that in a time Δt the energy which is removed from a radiation field of density $\rho(\nu)$ by the positive absorption of quanta by N oscillators all in $(n-1)$ quantum states is

$$\Delta E_{(n-1)\to n} = \tfrac{1}{3}n\pi e^2m^{-1}\, N\rho(\nu)\Delta t\,. \tag{13}$$

On the other hand it is well known that according to the classical theory the average rate at which a linear oscillator absorbs energy in a field of radiation is independent of the amplitude and is given[12] by $\frac{1}{3}\pi e^2m^{-1}\rho(\nu)$. Consequently in the interval Δt, N oscillators should absorb the energy

$$\Delta E_{abs} = \tfrac{1}{3}\pi e^2m^{-1}\, N\rho(\nu)\Delta t\,. \tag{14}$$

Because of the presence of the factor n, the quantum theory expression (13) for the positive absorption differs increasingly from (14) as the quantum number becomes larger. This discrepancy is not surprising, for we have not taken into account the fact that oscillators in the state $(n-1)$ may in the presence of a radiation field be induced to emit energy and pass to the state $(n-2)$ at a more rapid rate than when $\rho(\nu)=0$ and there are only spontaneous transitions.[13] Each of these excess or induced transitions may be thought of as returning to the ether or light wave the energy $h\nu_{st}$, a sort of regenerative effect. Hence we can take as the *net* absorption of energy in the time Δt by N oscillators in the state $(n-1)$ the expression

$$\Delta F = [B_{(n-1)\to n} - B_{(n-1)\to(n-2)}]\, h\nu\, N\rho(\nu)\, \Delta t\,.$$

[11] The result $\nu=\omega$ is obvious from an inspection of Eq. (3), as in the present case $W_n=nh\omega$, $\Delta n=\pm 1$.

[12] Cf. Planck, Wärmestrahlung, 4th Ed., Eqs. (260) and (159)

[13] If the electrons are in the orbit of lowest quantum number there can be only positive absorption because there are no still lower energy levels to which induced emission may take place. This case has been considered by Ladenburg (Zeit. f. Phys. **4**, 451, 1921); also by Ladenburg and Reiche (Naturwissenschaften, **27**, 584, 1923). These writers note the necessity of having the factor n in comparing the classical and positive quantum absorption.

This way of measuring absorption we shall term the differential rate of absorption in contrast to the true positive rate given by Eq. (13). Now by (12a) we have, on lowering the integer n by one unit, the result $B_{(n-1)\to(n-2)} = (n-1)\pi e^2/(3mh\nu)$, and hence $\Delta F = \Delta E_{abs}$. We thus see that in the limiting case of large quantum numbers, where Eq. (12a) is valid, the classical value (14) for the rate of absorption of energy is nothing but the differential rate of absorption in the quantum theory. This connection of the classical and quantum differential absorption we shall term the correspondence principle for absorption; it is a purely mathematical consequence of the correspondence principle for emission, which was used in deriving (12a). Another way of stating the results is that because of Eqs. (9) and (12a) the ratio $(B_{(n-1)\to n} - B_{(n-1)\to(n-2)})\,\rho(\nu)/A_{(n-1\to(n-2)}$ of the differential absorption to the spontaneous emission for any given orbit has for large n very nearly the classical value $c^3\rho(\nu)/(16\pi^3 m D_n^2 \nu^4)$ for the ratio of absorption to emission.[14]

5. Generalization of the Correspondence Principle for Absorption to an Arbitrary Non-degenerate Multiply Periodic System

We must now seek to examine whether an analogous correspondence principle for absorption holds for an arbitrary non-degenerate multiply periodic system, whose orbits are given by Eq. (1) and which represent essentially the most general type of motion amenable to present quantum theory methods. Here we can no longer enjoy such simplifications as the existence of a single quantum number and the results $\Delta n = \pm 1$, $\nu = \omega$ characteristic of a linear oscillator.

Let us consider a state s of quantum numbers n_1, n_2, n_3.[15] Then when a quantum of energy is absorbed by an orbit of this type the electron will pass to some other state of quantum numbers $n_1+\tau_1$, $n_2+\tau_2$, $n_3+\tau_3$.[16] Starting with the state s (not r) there is no emissive transition to a state t such that the frequency ν_{st} of the light thus emitted is just equal to the frequency ν_{rs} of that absorbed on passage to the state r. (This amounts

[14] Reference should be made while on the present subject to the discussion in section 158 of Planck's "Wärmestrahlung" (4th Ed.) ,which establishes an asymptotic connection between the classical and quantum formulas for statistical equilibrium. A correspondence principle for absorption could doubtless be derived from such considerations, but no explicit mention is made of the asymptotic connection of the classical absorption and the differential absorption for a single orbit (where thermodynamic equilibrium need not be assumed) which is the primary concern of the present paper.

[15] The generalization to cases of more than three quantum numbers is given in section 17.

[16] Not all the integers τ_1, τ_2, τ_3 need be positive but the state r must have a greater energy than s.

virtually to saying that in general energy levels are not evenly spaced, except, of course, for such an ideal case as the linear oscillator). However, if the state t has quantum numbers $n_1-\tau_1$, $n_2-\tau_2$, $n_3-\tau_3$, then for large values of n_1, n_2, n_3 the ratio ν_{rs}/ν_{st} approaches unity, for both these spectroscopic frequencies approach asymptotically the combination overtone $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$ of the orbital frequencies. Hence it is natural to define as the differential rate of absorption the excess of light absorbed by transitions from the state s to the state r over the energy returned to the ether by induced emissive transitions from the state s to the state t, where, as mentioned above, the differences between the quantum numbers for the states s and t are the same as the corresponding differences for the states r and s.

With this definition it follows that in the time Δt the differential absorption of light energy of frequency approximately $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$ by N atoms in the state s amounts to

$$\Delta F = \left[h\nu_{rs}\rho(\nu_{rs})B_{s\to r} - h\nu_{st}\rho(\nu_{st})B_{s\to t}\right] N\Delta t. \tag{15}$$

If we admit the validity of the correspondence principle for emission then Eqs. (7) and (9) show that for large quantum numbers this expression becomes approximately

$$\Delta F = \tfrac{2}{3}\pi^3 e^2 h^{-1}\left[\nu_{rs}\rho(\nu_{rs})\, D_r^{\,2} - \nu_{st}\rho(\nu_{st})\, D_s^{\,2}\right] N\Delta t$$

where for brevity we have written D_r for D^r (τ_1, τ_2, τ_3), etc. Also for large quantum numbers the discrete succession of quantum orbits becomes so nearly consecutive that differences may be replaced by differentials, just as is done in the derivation of the correspondence theorem for frequencies.[4] Consequently

$$\Delta F = \tfrac{2}{3}\pi^3 e^2 h^{-1}\left\{\rho(\nu) D_s^{\,2}(\nu_{rs}-\nu_{st}) + \nu_{st}\left(\frac{\partial\rho}{\partial\nu}\right) D_s^{\,2}(\nu_{rs}-\nu_{st}) + \nu_{st}\rho(\nu_{st})\delta D^2\right\} N\Delta t$$

where δ denotes the difference between an expression evaluated for the states s and t. It is, of course, to be understood that $\rho(\nu)$ and $\partial\rho/\partial\nu$ are to be evaluated with ν approximately equal to ν_{st} (or ν_{rs}). Now since we are dealing with the large quantum number limit, the correspondence theorem for frequencies tells us that ν_{st} may without sensible error be replaced by the combination overtone $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$ of the orbital frequencies, and similarly $\nu_{rs}-\nu_{st}$ is to the desired approximation equal to $\delta(\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3)$. Now since the system is non-degenerate, the amplitude D as well as the ω's and the energy W, is a function of the constants J_1, J_2, J_3 which according to the quantum conditions are equated to integral multiples of h, and consequently

$$\delta\omega_1 = \frac{\partial\omega_1}{\partial J_1}\delta J_1 + \frac{\partial\omega_1}{\partial J_2}\delta J_2 + \frac{\partial\omega_1}{\partial J_3}\delta J_3 = h\left(\tau_1\frac{\partial}{\partial J_1} + \tau_2\frac{\partial}{\partial J_2} + \tau_3\frac{\partial}{\partial J_3}\right)\omega_1, \text{ as } \delta J_k = \tau_k h\ .$$

Similar relations hold for $\delta\omega_2$, $\delta\omega_3$, and δD^2. The expression for the differential absorption therefore becomes

$$\Delta F = \tfrac{2}{3}\pi^3 e^2 \left[\rho(\omega_\tau) \left(\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \tau_3 \frac{\partial}{\partial J_3} \right) G_\tau + \frac{\partial \rho}{\partial \nu} G_\tau \left(\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \tau_3 \frac{\partial}{\partial J_3} \right) \omega_\tau \right] N \Delta t \qquad (16)$$

where we have written ω_τ for $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$ and G for $\omega_\tau[D(\tau_1, \tau_2, \tau_3)]^2$.

It is shown in part II of the present paper that Eq. (16) also gives the amount of energy which according to purely classical mechanics is absorbed in a time Δt by N systems, each with a multiply periodic orbit similar to the state s, when exposed to a radiation field of vanishing intensity except for frequencies in the vicinity of $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$. That is to say Eq. (16) also gives the part of the classical absorption due to resonance of the impressed waves with the combination overtone τ_1, τ_2, τ_3. The total classical absorption is, of course, the sum of resonance effects for all possible overtones.

We may therefore conclude that the correspondence principle for absorption holds even for an arbitrary multiply periodic orbit; i. e. assuming the validity of the correspondence principle for emission, the differential quantum absorption by a particular orbit approaches at high quantum numbers (usually long wave-lengths) the classical absorption of light of asymptotically corresponding frequency.[17]

[17] It is clearly to be understood that although in deriving (16) we have apparently utilized Eq. (9), the proof of (16) can be readily shown to be equally valid if (9) holds asymptotically, so that at ordinary wave-lengths $A_{\tau\to s}$ might, for instance, be determined by averaging the frequencies and amplitudes in the manner given in Eq. (10). In this connection it is interesting to note that it is not difficult to prove that for the special case of a linear oscillator the differential quantum absorption and the classical absorption are exactly equal to each other even for low quantum numbers provided $A_{\tau\to s}$ is determined by taking the + signs in (10), with the integer m arbitrary. On the other hand it was found in section 2 that to make the mean free time of spontaneous emission in the quantum theory just equal the classical time of transit from one state to another in the absence of external radiation (i.e. $\rho(\nu)=0$), it is necessary to take the − signs in (10) and $m=-3$.

Another way of formulating the correspondence principle for absorption is as follows. The quantity $h\nu A_{\tau\to s}$ plays in the quantum theory the rôle of a coefficient of emission, and may be denoted by a. Using arguments similar to those employed in the derivation of (16), it is easy to show that the differential absorption has for large quantum numbers very approximately the value

$$\Delta F = (c^3/8\pi) \left[\rho(\omega_\tau) \left(\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \tau_3 \frac{\partial}{\partial J_3} \right) (a\omega_\tau^{-3}) + a\omega_\tau^{-3} \frac{\partial \rho}{\partial \nu} \left(\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \tau_3 \frac{\partial}{\partial J_3} \right) \omega_\tau \right] N \Delta t .$$

In the present paper we have proved the correspondence principle for absorption after assuming the validity of the ordinary correspondence principle for emission. The procedure could equally well have been reversed. That is to say, assuming an asymptotic connection of the differential absorption in the quantum theory with the corresponding absorption in the classical mechanism we could show that the spontaneous emission in the quantum theory must approach equality with the classical formula (8) for the spontaneous radiation of energy.

The classical formula (8) for spontaneous emission involves rather more electrodynamics than the analogous classical formula for absorption, given in Eq. (16), for the derivation of the former is based on retarded potentials and the validity of the field equations, while the latter requires simply that there be a certain incoherence in the radiation field and that the change in the energy be equal to the scalar product of the impressed force and the velocity of the particle, a very broad mechanical principle. As pointed out in the preceding paragraph, the asymptotic values of the probability coefficients $A_{\tau\to s}$, $B_{\tau\to s}$, etc., can be derived from the absorption instead of the emission viewpoint. Because of this alternative the correspondence principle for evaluating these coefficients is in a certain sense less electrodynamical and more purely mechanical in nature than it has sometimes been considered to be. From this standpoint the connection between the quantum theory and classical electrodynamics comes through the assumption of Wien's law in obtaining Eq. (7), for in Einstein's derivation of the Planck radiation formula considerations of statistical equilibrium without resorting to Wien's law do not suffice to determine $B_{\tau\to s}/A_{\tau\to s}$. If we assumed both the correspondence principle for emission and that for absorption it would be possible to dispense with the need of assuming (rather than proving) Wien's law. Virtually this procedure was adopted by Planck, who studied the special cases of the linear oscillator and rotating dipole.[18] As pointed out to the writer by Prof. Ehrenfest, the content of the present paper may be regarded as placing such a procedure on a more general basis.[19]

This is, however, nothing but the classical theory formula for the coefficient of absorption in terms of the coefficient of emission as can be seen from (8) and (16). This way of stating the correspondence principle for absorption does not utilize (9) even asymptotically.

[18] Wärmestrahlung, 4th Ed., section 158. Cf. footnote [14].

[19] The writer has just learned (Sept. 1924) that in unpublished computations made at Copenhagen, J. C. Slater has independently derived an absorption formula similar to Eq. (16) and has also noted the asymtotic connection of the classical and quantum absorption discussed in sections 4 and 5.

6. The General Correspondence Principle Basis for Kramers' Dispersion Formula

In a recent note[20] H. A. Kramers has proposed a formula for dispersion which is a modification of an equation previously developed by Ladenburg and Reiche,[21] and which must be regarded as a distinct advance in the problem of reconciling dispersion with quantum phenomena. The physical principle underlying Kramer's formula is that of "virtual" or "ghost" resonators, first suggested by Slater[22] and elaborated by Bohr, Kramers, and Slater.[23] According to this viewpoint the dispersion is not to be calculated by considering the actual orbit (stationary state) as reacting classically to the impressed waves. Instead, the stationary states appear to be unaffected except for occasional quantum leaps, but the dispersion is to be computed as due to a set of hypothetical linear oscillators whose frequencies are the spectroscopic ones rather than those of the orbits. The introduction of these virtual resonators is, to be sure, in some ways very artificial, but is nevertheless apparently the most satisfactory way of combining the elements of truth in both the classical and quantum theories. In particular this avoids the otherwise almost insuperable difficulty that it is the spectroscopic rather than the orbital frequencies, i.e. ν rather than $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$, which figure in dispersion.[24]

Kramer's formula for the polarization due to N_s electrons in the state s is[25]

$$P_s = \frac{e^3}{32\pi^4}\left[\sum_r \frac{A_{r\to s}}{\nu_{rs}^2(\nu_{rs}^2-\nu^2)} - \sum_t \frac{A_{s\to t}}{\nu_{st}^2(\nu_{st}^2-\nu^2)}\right] N_s\mathscr{E} \ . \quad (17)$$

[20] H. A. Kramers Nature, May 10, 1924, p. 673; Aug. 30, 1924, p. 310.

[21] Ladenburg and Reiche, Naturwissenschaften **27**, 584 (July 6, 1923). Ladenburg and Reiche's formula differs from that of Kramers, given in (17), principally in that the negative terms are not included.

[22] Slater, Nature, Mar. 1, 1924, p. 307

[24] Bohr, Kramers and Slater, Phil. Mag. **47**, 785 (1924)

[24] Reference should, however, be made to the alternative explanation proposed by Darwin (Proc. Nat. Acad. **9**, 25, 1923).

[25] This equation differs by a factor 1/3 from the formula in Kramer's note to Nature, as (17) is written for the case where all orientations of the atom are equally probable, while Kramers' original equation is for the case that the atoms are always so oriented that their free vibrations are parallel to the impressed electric field. The assumption that all orientations are equally probable may appear at first sight contrary to the assumption that the system is non-degenerate, as all three intrinsic frequencies (the criterion for non-degeneracy) can appear only when the atomic force field does not have a spacial symmetry, and with such a dissymmetry the quantum conditions allow only certain particular orientations. However, we may imagine the asymmetrical internal atomic force fields to have all possible orientations in the different atoms. This artifice,

where $\mathcal{E}$ is the electric intensity and ν is the frequency of the impressed wave. The summation with respect to r is to be taken over all states of energy content higher than s, while that for t over all orbits of lower energy content than s. Each term in this equation represents the dispersion due to a "virtual oscillator" and corresponds to a particular quantum transition starting from the state s. Each oscillator in the t summation gives a negative dispersion, corresponding to Einstein's negative absorption, mentioned in section 3.

Kramers states that (17) merges asymptotically for large quantum numbers into the classical formula for the polarization. It is our purpose to show that this is true not just when the quantized system is a linear oscillator, but also when it is the most general type of non-degenerate multiply periodic orbit.[26] Just as in section 5, the upward transitions ($s \to r$) and the downward ones ($s \to t$) may be mated together in such a way that in each pair ν_{rs}/ν_{st} converges to unity for large quantum numbers. This is accomplished, just as before, by making the differences in quantum numbers between the states r and s the same as those between the states s and t, and we shall denote these differences by τ_1, τ_2, τ_3. Each pair contributes a positive and a negative term in (17) and their net effect is a "differential dispersion" very analogous to the differential absorption mentioned in sections 4 and 5. In precisely the same way that (16) is

---

introduced for simplicity, enables one to have a random spacial distribution of orbits without having degenerate systems.

In this connection it is interesting to note that if the axes of the asymmetrical atomic force fields had the same orientations in all atoms, as is the case when the dissymetry is due to a constant external magnetic or electric field, then we should expect the virtual resonators to have only certain particular orientations, as only a few quantized spacial positions can be assumed by the atoms. By analogy with the classical theory results mentioned after Eq. (41) in section 15 of part II, there might then be possible a polarization in the y and z directions even though the electric intensity coincides with the x axis. The polarization and electric vectors would then have different directions, as in a crystal. Therefore the application of a weak constant external field would cause an abrupt dissymmetry in the index of refraction just as it creates an abrupt outstanding polarization of the spontaneously radiated light (cf. N. Bohr, the Fundamental Postulates of the Quantum Theory, Supplement of Proc. Cambr. Phil. Soc., p. 27). These sudden discontinuities are not found in the classical theory because the spacial distributions are continuous rather than quantized.

[26] In his first note Kramers did not give his proof of the desired asymptotic connection. Inasmuch as the formula for the dispersion by an arbitrary multiply periodic orbit, to be developed in section 15 of part II, did not appear to have previously been given and was not mentioned in Kramers' note, the writer supposed that Kramers' prior demonstration was somewhat less general than that of the present paper. Since writing this paper, the author has learned, however, that Kramers' unpublished computations are of the same scope and generality as those in § 15. Cf. Kramers' second note[20].

derived from (15) we find that for very large quantum numbers the contribution of one particular pair to the polarization becomes

$$P_s = \frac{1}{6}e^2 \left[\frac{1}{(\omega_\tau^2 - \nu^2)} \left(\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \tau_3 \frac{\partial}{\partial J_3} \right) G_\tau - \right.$$
$$\left. \frac{2\omega_\tau G_\tau}{(\omega_\tau^2 - \nu^2)^2} \left(\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_1} + \tau_3 \frac{\partial}{\partial J_2} \right) \omega_\tau \right] N_s \mathscr{E} \quad (18)$$

where G_τ and ω_τ are defined as in Eq. (16).

It is proved in section 15 that (18) also gives the polarization which according to the classical theory would be produced by resonance between the impressed frequency and the combination overtone ω_τ in a multiply periodic orbit of the same size and shape as the state s. The differential dispersion in Kramers' theory therefore approaches asymptotically the classical dispersion for the actual orbit. The asymptotic connection is thus very similar to that previously encountered in the correspondence principle for absorption.

It is particularly interesting to note that although both the positive and negative terms in the differential dispersion taken separately represent a type of dispersion characteristic of a linear oscillator (except that oscillators with negative dispersion correspond to no ordinary physical reality, as they would have to possess a negative mass) the difference of the two terms approaches asymptotically a more complicated type of dispersion appropriate to the general multiply periodic orbit. For Eq. (18) contains a term in $(\omega_\tau^2 - \nu^2)^{-2}$ as well as the familiar term in $(\omega_\tau^2 - \nu^2)^{-1}$ characteristic of the ideal resonator. The fact that the dispersion for the simple virtual oscillators thus merges asymptotically into the more complex classical dispersion for the actual orbits must be regarded as an important argument for the virtual resonator viewpoint, as such a connection is not an obvious outcome of the theory. Especially does this asymptotic connection increase our faith in the particular form of virtual oscillator theory embodied in Kramers' dispersion formula, and makes it easier to accept the rather artificial negative dispersion terms thus involved.[27]

[27] It is to be noted that if $A_{\tau \to s}$ is determined in accordance with Eq. (10), the + signs being taken, then the differential dispersion in the Kramers' formula is for the special case of a quantized linear oscillator easily proved identical even at low quantum numbers with the classical dispersion. It was pointed out in footnote [17] that this kind of formula also makes the differential and classical absorption identical for the ideal oscillator without passing to the high quantum number limit. These facts taken together furnish a limited amount of evidence for the use of the average of the form (10) (with + signs) rather than the alternative logarithmic average which Hoyt has shown explains observed intensities equally well in the instances studied by him.[6]

THE ABSORPTION OF RADIATION BY MULTIPLY PERIODIC ORBITS, AND ITS RELATION TO THE CORRESPONDENCE PRINCIPLE AND THE RAYLEIGH-JEANS LAW

PART II. CALCULATION OF ABSORPTION BY MULTIPLY PERIODIC ORBITS

By J. H. Van Vleck

Abstract

This part is largely dynamical and presents the mathematical analysis on which part I is largely based. It gives the calculation of the absorption by a multiply periodic orbit in a field of radiation and the corresponding perturbations in the orbits, also of the polarization produced by incident waves, causing dispersion, and the distortion of the orbit by its own radiation. Most of these calculations have previously been given only for special cases.

7. Introduction. General Dynamical Theory of Multiply Periodic Orbits

It is the purpose of part II primarily to calculate according to the classical mechanics the average absorption and certain other properties of a multiply periodic orbit in a radiation field. Besides being of some interest taken by itself, this computation is necessary to furnish a mathematical basis for much of the material in parts I and III. Although multiply periodic orbits have been extensively studied by many investigators, the calculation of the absorption appears to have hitherto[19] been published for only a few special cases, such as the linear oscillator and one-dimensional rotating dipole whose rates of absorption have been examined by Planck and Fokker respectively.[28]

Let us consider an electron whose Cartesian coordinates are x, y, z, and which in the absence of a radiation field is subject to forces which can be derived from a static potential $V(x, y, z)$. We thus assume that the system has only three degrees of freedom and hence but three intrinsic frequencies. These restrictions are made only in order to avoid excessively complex notation. (The appropriate modifications for the more general case of several electrons and hence more than 3 degrees of freedom are given in section 17.) Instead of using the ordinary three second order

[28] M. Planck, Wärmestrahlung, Part IV, Ch. 1, 4th Ed. (Part IV, Ch. 2 of Engl trans. by Masius). A. Fokker, Ann. der Phys. **43**, 812 (1914).

Newtonian equations ($m\ddot{x} = -\partial V/\partial x$, etc.), the motion can equally well be specified by the six first order Hamiltonian canonical equations

$$\frac{dp_x}{dt} = -\frac{\partial H}{\partial x}, \qquad \frac{dx}{dt} = \frac{\partial H}{\partial p_x}, \text{ etc.} \tag{19}$$

where the Hamiltonian function H is the total energy $T+V$, so that

$$H = (1/2m)\,(p_x^2 + p_y^2 + p_z^2) + V(x, y, z). \tag{20}$$

It is convenient to make a change of variables from x, y, z, p_x, p_y, p_z to new coordinates $w_1, w_2, w_3, J_1, J_2, J_3$. In a multiply periodic system it is possible to find a transformation possessing the three following properties.

(I) The six original coordinates and momenta, x, p_x, etc. are periodic in each of the w's with modulus 2π, and so can be expanded in series of the form[29]

$$x = \Sigma_{\tau_1\tau_2\tau_3} A(\tau_1,\tau_2,\tau_3) \exp\,[2\pi i(\tau_1 w_1 + \tau_2 w_2 + \tau_3 w_3)] \tag{21}$$

with similar equations for y, z, p_x, p_y, p_z, except that different coefficients appear in front of the exponentials. The summation is to be extended over all positive and negative integral values of τ_1, τ_2, τ_3. By Fourier's theorem

$$A(\tau_1,\tau_2,\tau_3) = (2\pi)^{-3}\int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} x \exp[-2\pi i(\tau_1 w_1+\tau_2 w_2+\tau_3 w_3)]\times dw_1 dw_2 dw_3 \tag{22}$$

and consequently $A(-\tau_1,-\tau_2,-\tau_3)$ is the conjugate of $A(\tau_1,\tau_2,\tau_3)$, making the resulting series real despite the imaginary exponentials. The use of this complex variable notation makes the equations very much more compact and easier to handle than if sines and cosines are employed.

(II) The expression $p_x dx + p_y dy + p_z dz - J_1 dw_1 - J_2 dw_2 - J_3 dw_3$ is the exact differential dF of some function $F(x, y, z, w_1, w_2, w_3)$ when the p's and J's are expressed in terms of x, y, z, w_1, w_2, w_3 by means of the transformation formulas. Such a transformation is often termed a contact transformation. We shall furthermore suppose the time average of the function F is zero, so that the expression $p_x\dot{x} + p_y\dot{y} + p_z\dot{z}$ (which is twice the kinetic energy) is on the average equal to $J_1\dot{w}_1 + J_2\dot{w}_2 + J_3\dot{w}_3$.

(III) The energy or Hamiltonian function H given in (20) is expressible in terms of J_1, J_2, J_3 alone, (i.e., the w's drop out when H is expressed in terms of the J's and w's by means of the transformation formulas (21)).

In virtue of the fundamental theorem that a contact transformation preserves the Hamiltonian form,[30] the differential equations for determining how the w's and J's vary with the time may be written

[29] For simplicity in printing we have denoted the exponential function $e^{i\theta}$ by $\exp[i\theta]$.

[30] For proof of this theorem, see, for instance, appendix 4 of Sommerfeld's Atombau.

$$\frac{dJ_k}{dt} = -\frac{\partial H}{\partial w_k}, \qquad \frac{dw_k}{dt} = \frac{\partial H}{\partial J_k} \tag{23}$$

where the subscript k may have the values 1, 2, or 3. This range of values for k is to be understood throughout, while the letter i will be reserved strictly for $\sqrt{-1}$. Now by (III) we have $\partial H/\partial w_k = 0$, and hence the J's are constants independent of the time. Therefore the time derivatives of w's, which by (23) are functions only of the J's (since H involves only the J's), are also constant, and consequently

$$w_k = \omega_k t + \epsilon_k \tag{24}$$

where the ϵ_k are arbitrary constants or phase angles, while the expressions $\omega_1, \omega_2, \omega_3$ are functions only of the J's and are the intrinsic frequencies of the system. We shall suppose that the system is non-degenerate, so that there is no commensurability relation of the form $l\omega_1 + m\omega_2 + n\omega_3 = 0$, where l, m, and n are integers.[31]

By introducing the values of the w's given in (24) and using the conjugate property mentioned in connection with Eq. (22), it is seen that Eq. (21) may be replaced by a series of cosine terms of the form (1) provided

$$[X(\tau_1,\tau_2,\tau_3)]^2 = 4A(\tau_1,\tau_2,\tau_3)\, A(-\tau_1, -\tau_2, -\tau_3)\,, \tag{25}$$

$$\gamma_{\tau_1\tau_2\tau_3}{}^{(x)} = 2\pi(\tau_1\epsilon_1 + \tau_2\epsilon_2 + \tau_3\epsilon_3) + \operatorname{arc} A(\tau_1,\tau_2,\tau_3).$$

According to the quantum conditions, the constants J_1, J_2, J_3 should be equated to integral multiples of h, as it follows from (23), (24), and the latter part of postulate (II), that the J's have the properties demanded in Eq. (2).[32] However, this need not concern us immediately, as the present computation is based on purely classical mechanics without quantum modifications.

8. Derivation of the Equations for the Perturbations Caused by Adding a Radiation Field

In the preceding paragraphs we have reviewed briefly some of the general properties of multiply periodic dynamical systems which it will be necessary to use in the forthcoming analysis.[33] Let us now assume

[31] When there exists a commensurability relation of the form given above, it is possible to eliminate one of the three intrinsic frequencies by making one of the ω's zero through a transformation of coordinates. Cf. P. S. Epstein, Ann. der Phys. **51**, 179 (1916).

[32] In writing Eq. (2) W has been used in place of H, as these two letters can be used interchangeably because the Hamilton function defined in (20) is identical with the energy. The factor 4 in Eq. (25) is due to the fact that the number of terms in the series (1) is half that in (21) because of the requirement in (1) that $\tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3$ be positive.

[33] For greater detail concerning the transformation theory outlined in §7 see Appendix 4 of Sommerfeld's Atombau or Ch. XI of Whittaker's Analytical Dynamics.

that a field of radiation is present. The system will then no longer be conservative and our problem is to compute the rate at which work is done by the forces exerted by the light waves.

As is usual in the study of analogous problems, we shall commence by assuming that the incident radiation consists entirely of monochromatic waves of frequency ν, and that the y and z components of the electric intensity are zero, while the peak value of the x component will be denoted by E_x. It is clear that the Newtonian differential equations now become

$$m\ddot{x} = -\partial V/\partial x - eE_x\cos[2\pi(\nu t+\epsilon_0)]; \quad m\ddot{y} = -\partial V/\partial y; \quad m\ddot{z} = -\partial V/\partial z; \quad (25a)$$

where e is the absolute value of the electronic charge and ϵ_0 is an arbitrary phase angle.

If we were to study the perturbations in the motion produced by the incident wave purely with the aid of Eqs. (25a) it would be impossible to make further progress without specializing the form of the potential function V, as was done by Planck and Fokker. However, it is quite a different story when we seek to compute the perturbations not in x, y, z but in the "angle variables" w_1, w_2, w_3 and their conjugate momenta J_1, J_2, J_3, whose properties were sketched in section 7. Because of their intimate connection with the orbital frequencies ω_k, these coordinates possess important intrinsic properties. In fact by using them rather than x, y, z, which is the essential feature of the present calculation, the periodic properties of the system come to light even without knowing the form of V, and it is thus possible to isolate the resonance between the frequency of the impressed wave and each combination overtone $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$ of the orbital vibrations. This resonance will be shown to depend, in the average, only on the associated amplitude $X(\tau_1, \tau_2, \tau_3)$.

To get the equations for the perturbations of the J's and w's, we need simply note that the complete x component of force in Eq. (25a) may be written as the negative x derivative of the function $V+xeE_x\cos[2\pi(\nu t+\epsilon_0)]$. Consequently the equations of motion in Cartesian coordinates may be thrown into the canonical form (19) provided in these equations we replace H by a new Hamiltonian function H' which is no longer the sum of the kinetic and potential energies but instead has the value

$$H' = H + xeE_x\cos[2\pi(\nu t + \epsilon_0)] \quad (25b)$$

where H is the Hamiltonian function in the absence of the radiation field, which is equal to the energy and given by (20). Since by postulate (II), the transformation from x, y, z, p_x, p_y, p_z to the w's and J's is a contact transformation, and since it is a fundamental theorem that such a transformation preserves the Hamiltonian form even though

the Hamiltonian function contains the time explicitly, the equations for the w's and J's still have the canonical form (23). We must, however, utilize in place of H the new Hamiltonian function H' defined above, but expressed now in terms of the w's and J's rather than Cartesian coordinates and momenta. We therefore have

$$\frac{dw_k}{dt} = \frac{\partial H}{\partial J_k} + \frac{\partial x}{\partial J_k} eE_x \cos[2\pi(\nu t + \epsilon_0)];$$

$$\frac{dJ_k}{dt} = -\frac{\partial H}{\partial w_k} - \frac{\partial x}{\partial w_k} eE_x \cos[2\pi(\nu t + \epsilon_0)].$$

Now by postulate (III) concerning the transformation formulas, H is a function only of the J's, so that $\partial H/\partial w_k = 0$. Consequently using the value for x given in (21), and writing $\cos[2\pi(\nu t+\epsilon_0)]$ as the sum of two imaginary exponentials, we get

$$\frac{dw_k}{dt} = \frac{\partial H}{\partial J_k} + \tfrac{1}{2} eE_x \sum \left[\frac{\partial A(\tau_1, \tau_2, \tau_3)}{\partial J_k} \left\{ \exp[2\pi i(\tau_1 w_1 + \tau_2 w_2 + \tau_3 w_3 + \nu t + \epsilon_0)] \right.\right.$$
$$\left.\left. + \exp[2\pi i(\tau_1 w_1 + \tau_2 w_2 + \tau_3 w_3 - \nu t - \epsilon_0)] \right\} \right]; \quad (26)$$

$$\frac{dJ_k}{dt} = -\pi i e E_x \sum \left[\tau_k A(\tau_1, \tau_2, \tau_3) \left\{ \exp[2\pi i(\tau_1 w_1 + \tau_2 w_2 + \tau_3 w_3 + \nu t + \epsilon_0)] \right.\right.$$
$$\left.\left. + \exp[2\pi i(\tau_1 w_1 + \tau_2 w_2 + \tau_3 w_3 - \nu t - \epsilon_0)] \right\} \right]. \quad (27)$$

9. Power Series Integration of Equations for Perturbations in w's and J's

If no impressed electric field were present, the J's would have constant values, which we shall denote by J_1^0, J_2^0, J_3^0, while the w's would be given by $\omega_k t + \epsilon_k$ (cf. Eq. 24). Our problem is essentially to compute the departures of the w's and J's from these values. This can be done by the standard method of determining the perturbations of the w's and J's as power series in the constant E_x which represents the amplitude of the expressed wave, and which may be assumed to be small. For the results which we shall desire, it will only be necessary to determine the perturbations in the J's to second powers of E_x, and those in the w's merely to the first power.

To determine the perturbations in the J's to first powers of E_x we can substitute in the right hand side of Eq. (27) the unperturbed values $\omega_k t+\epsilon_k$ for the coordinates w_k, and in this equation we may similarly regard the amplitudes $A(\tau_1,\tau_2,\tau_3)$ as having the constant values $A^0(\tau_1,\tau_2,\tau_3)$ which result when their arguments J_1, J_2, J_3 are equated to J_1^0, J_2^0, J_3^0. Eq. (27) thus becomes

$$dJ_k/dt = -\pi ieE_x{\sum}_\tau{}'\left\{\tau_k A_\tau \exp[2\pi i(\epsilon_\tau-\epsilon_0)]\exp[2\pi i(\omega_\tau t-\nu t)]\right\} \quad (28)$$

where for brevity we have written ω_τ for $\tau_1\omega_1+\tau_2\omega_2+\tau_3\omega_3$, ϵ_τ for $\tau_1\epsilon_1+\tau_2\epsilon_2+\tau_3\epsilon_3$ and A_τ for $A^0(\tau_1,\tau_2,\tau_3)$. Σ_τ denotes the summation with respect to τ_1, τ_2, τ_3, and it is to be clearly understood that this τ summation is a triply infinite summation wherever it appears, in distinction from the k, l summations later introduced which involve only three terms. In (28) terms are wanting in $(\omega_\tau t+\nu t)$ and $(\epsilon_\tau+\epsilon_0)$, but instead of writing these terms explicitly we have inserted a prime above the summation sign to indicate that the entire summation is to be repeated with $-\nu$ and $-\epsilon_0$ everywhere replaced by $+\nu$ and $+\epsilon_0$. The symbol Σ' is used with this significance throughout, in order to halve the number of terms which must be written explicitly.

We shall suppose the field is applied at $t=0$, so that at this instant the w's and J's have their unperturbed values. Taking account of these initial conditions we have on integration of (28)

$$\Delta J_k=\tfrac{1}{2}eE_x{\sum}_\tau{}'\left\{\tau_k A_\tau\exp[2\pi i(\epsilon_\tau-\epsilon_0)]\,\frac{1-\exp[2\pi i(\omega_\tau t-\nu t)]}{\omega_\tau-\nu}\right\} \quad (29)$$

where $\Delta J_k=J_k-J_k{}^0$.

Having just found the perturbations in the J's to first order terms in E_x we are now in a position to determine the w's to this degree of approximation. We first observe that because of the fluctuations in the J's, the expressions $\partial H/\partial J_k$ appearing in (26) do not have their original unperturbed values ω_k, but may, as they are functions only of the J's, be expanded as power series in ΔJ_1, ΔJ_2, ΔJ_3. We may therefore, to the order of approximation under consideration, replace $\partial H/\partial J_k$ by $\omega_k+\Sigma_{(l=1,2,3)}(\partial\omega_k/\partial J_l)\Delta J_l$ where $\partial\omega_k/\partial J_l$ denotes $\partial^2 H/\partial J_k\partial J_l$ evaluated with each J_k equal to $J_k{}^0$. Similarly we may in the right hand side of Eq. (26) set $w_k=\omega_k t+\epsilon_k$ and replace each expression $\partial A(\tau_1,\tau_2,\tau_3)/\partial J_k$ by its unperturbed value, which will be denoted by $\partial A_\tau/\partial J_k$. When these sub-

stitutions are made, and account is taken of the initial conditions, the integration of (26) gives

$$\Delta w_k = eE_x \sum_\tau{}' \left\{ \left(\Sigma_{(l=1,2,3)} \tau_l \frac{\partial \omega_k}{\partial J_l} \right) A_\tau \exp[2\pi i(\epsilon_\tau - \epsilon_0)] \times \right.$$
$$\frac{2\pi(\omega_\tau - \nu)t - i\{1 - \exp[2\pi i(\omega_\tau - \nu)t]\}}{4\pi(\omega_\tau - \nu)^2} +$$
$$\left. i \frac{\partial A_\tau}{\partial J_k} \exp[2\pi i(\epsilon_\tau - \epsilon_0)] \frac{1 - \exp[2\pi i(\omega_\tau - \nu)t]}{4\pi(\omega_\tau - \nu)} \right\} \quad (30)$$

where $\Delta w_k = w_k - \omega_k t - \epsilon_k$.

10. Computation of the Perturbations ΔJ_k to 2nd Order terms E_x

We can now commence the calculation of the J_k to second powers of E_x. We take the right hand side of (27) and expand the amplitudes as power series in the perturbations ΔJ_k, and similarly develop the factor $\exp[2\pi i(\tau_1 w_1 + \tau_2 w_2 + \tau_3 w_3)]$ as series in the Δw's. We thus get to the approximation under consideration

$$\frac{dJ_k}{dt} = -i\pi e E_x \, \Sigma_\tau' \, \Sigma_{(l=1,2,3)} \left\{ \tau_k \left(A_\tau + 2\pi i \tau_l \Delta w_l A_\tau + \frac{\partial A_\tau}{\partial J_l} \Delta J_l \right) \times \right.$$
$$\left. \exp[2\pi i(\epsilon_\tau - \epsilon_0)] \exp[2\pi i(\omega_\tau - \nu)t] \right\}. \quad (31)$$

The values of the ΔJ's and Δw's given by (29) and (30) can now be substituted. We have then to deal with the product of two triply infinite series, which we shall assume can be multiplied together term by term. The only significant terms for our purposes are those which arise from the product of a term in τ_1, τ_2, τ_3 into one in $-\tau_1$, $-\tau_2$, $-\tau_3$. Other terms would involve a factor of the form $\exp[2\pi i(\epsilon_\tau + \epsilon_{\tau'})]$ and so would vanish when we average over an ensemble of systems having the same initial values of the J's but haphazard values of the phase angles ϵ_1, ϵ_2, ϵ_3. However, taking this average does not appear really necessary to make the other terms mentioned above disappear, as it can be shown that such terms vanish if we assume that instead of random waves applied only at $t=0$ there are a large number of waves of the same frequency whose times of application are haphazard.[34]

[34] If the integration over all values of ϵ_1, ϵ_2, ϵ_3 is omitted (but the integration over ϵ_0 is retained) there would appear in Eq. (34) "Poisson terms" of the type form $R_\rho(\omega_\tau)t\cos[2\pi(\omega_\tau + \omega_{\tau'})t + \gamma]$. This is on the assumption that the field is applied at $t=0$. If instead the field were applied at $t=-T$, the change in a Poisson term from time 0 to

Furthermore in multiplying a term in (τ_1,τ_2,τ_3) by one in $(-\tau_1,-\tau_2,-\tau_3)$, we may omit such products as involve the factor $\exp[\pm 4\pi i\epsilon_0]$, as these vanish when we average over a large number of impressed waves whose phase angles ϵ_0 are haphazard.[35] The first order terms in E_x also vanish when such an average is taken. We thus get for the average value of dJ_k/dt.

$$\overline{\frac{dJ_k}{dt}}=\tfrac{1}{2}\pi i e^2E_x^2\ \Sigma_\tau'\Bigg\{\tau_kA_\tau A_{-\tau}\Bigg(i\Sigma_{(l=1,2,3)}\tau_l\frac{\partial\omega_\tau}{\partial J_l}\Bigg)\times \frac{i\{1-\exp[2\pi i(\omega_\tau-\nu)t]\}-2\pi(\omega_\tau-\nu)t\exp[2\pi i(\omega_\tau-\nu)t]}{(\omega_\tau-\nu)^2}+ \tau_k\Sigma_{(l=1,2,3)}\Bigg(\tau_lA_\tau\frac{\partial A_{-\tau}}{\partial J_l}+\tau_lA_{-\tau}\frac{\partial A_\tau}{\partial J_l}\Bigg)\frac{1-\exp[2\pi i(\omega_\tau-\nu)t]}{\omega_\tau-\nu}\Bigg\}$$

where $A_{-\tau}$ denotes $A(-\tau_1,-\tau_2,-\tau_3)$. By combining terms in $(\omega_\tau\mp\nu)$ with those in $-\omega_\tau\pm\nu$, this expansion can be written as the sum of only real terms. (In doing this one must remember to reverse the signs of τ_k and τ_l as well as ω_τ.) Using the formula for X_τ^2 given in (25) and performing the time integration we thus get

$$\overline{\Delta J}_k=\tfrac{1}{8}e^2E_x^2\Sigma_\tau'\Bigg\{\tau_kX_\tau^2\Bigg(\Sigma_{(l=1,2,3)}\tau_l\frac{\partial\omega_\tau}{\partial J_l}\Bigg)\times \frac{[-2+2\cos2\pi(\omega_\tau-\nu)t]+2\pi(\omega_\tau-\nu)t\sin2\pi(\omega_\tau-\nu)t}{(\omega_\tau-\nu)^3}+ \tau_k\Bigg(\Sigma_{(l=1,2,3)}\tau_l\frac{\partial X_\tau^2}{\partial J_l}\Bigg)\frac{1-\cos2\pi(\omega_\tau-\nu)t}{(\omega_\tau-\nu)^2}\Bigg\}\quad(32)$$

In this equation the summation Σ_τ' is to be taken not over *all* positive and negative values of τ_1, τ_2, τ_3, but over all such values subject to the

t would be $R\rho(\omega_\tau)\,(t+T)\cos[2\pi(\omega_\tau+\omega_{\tau'})(t+T)+\gamma]-R\rho(\omega_\tau)T\cos[2\pi(\omega_\tau+\omega_{\tau'})T+\gamma]$ and so practically vanishes on the average when we allow the values of T to be uniformly distributed in an interval large compared to a period. Besides these Poisson terms, there would also be trigonometric and constant terms, but their effect during a long time interval is unimportant compared to the "secular" or linear variation with the time given in (34). It is to be clearly understood that integrating over the time of application $-T$ is quite different from integrating over the phase angle ϵ_0 of the impressed wave, and the former integration is required over and above the latter in order to make the cross terms vanish without averaging over ϵ_1, ϵ_2, ϵ_3.

[35] The remarks concerning the vanishing of products may be summarized in the statement that the only significant terms result from the multiplication of a term $\omega_\tau\mp\nu$ in (31) by a term $-\omega_\tau\pm\nu$ in Δw or ΔJ. In forming a term in $-\omega_\tau\pm\nu$ in Eqs. (29) and (30) care must be taken to reverse the sign of τ_k (in 29) or τ_l (in 30), as well as the signs of ω_τ and ν. The sign of ω_k in (30), however, should not be reversed, as ω_k does not involve the τ's.

restriction that ω_τ be positive. This normalization, which halves the number of terms in the summation, is to be understood throughout the article whenever the equations involve the real amplitudes X_τ, D_τ, etc. instead of the complex coefficients A_τ.

11. Passage from Monochromatic to Continuous Radiation

Hitherto we have assumed that the impressed wave is strictly monochromatic. Actually, however, we must deal with an aggregate of waves having a continuous distribution of frequencies. In our calculations the passage from the monochromatic to the continuous case will be almost identical with the procedure adopted by Planck in his computations for the absorption by an oscillator, and for a more detailed or rigorous justification for the various steps the reader is referred to Planck's original work.[28]

If $\rho(\nu)$ denote the specific energy density of radiation, the amount of energy between frequencies ν and $\nu+d\nu$ is $\rho(\nu)d\nu$. On the other hand it is well known that in a homogeneous radiation field, each of the six components of the electric and magnetic intensity contribute equally to the energy, so that the mean value of $(\mathscr{E}^2+H^2)/8\pi$, the energy in the medium, is $3\,\bar{\mathscr{E}}_x^2/4\pi$. We thus see that for the part $\mathscr{E}_x$ of the electric intensity due to vibrations in the frequency range ν, $\nu+d\nu$, we have $\bar{\mathscr{E}}_x^2=\frac{4}{3}\pi\rho(\nu)d\nu$. Now in our problem $\mathscr{E}_x=E_x\cos 2\pi(\nu t+\epsilon_0)$, making $\bar{\mathscr{E}}_x^2=\frac{1}{2}E_x^2$. Therefore we see that E_x^2 is to be identified with the element $\frac{8}{3}\pi\rho(\nu)d\nu$.

Our problem is now to replace E_x^2 by $\frac{8}{3}\pi\rho(\nu)d\nu$ in Eq. (32) and integrate over all values of the frequency ν of the impressed light wave. If x denote the expression $(\nu-\omega_\tau)t$, then it is clear that the contribution to ΔJ of one particular term τ_1, τ_2, τ_3 in the multiple Fourier expansion may be written

$$(\overline{\Delta J_k})_\tau=\tfrac{1}{3}\pi e^2\tau_k X_\tau^2\left(\Sigma_{(l=1,2,3)}\tau_l\frac{\partial\omega_\tau}{\partial J_l}\right)t^2\int\frac{[2-2\cos(2\pi x)-2\pi x\sin(2\pi x)]}{x^3}\rho(\nu)dx$$

$$+\tfrac{1}{3}\pi e^2\tau_k\left(\Sigma_{(l=1,2,3)}\tau_l\frac{\partial X_\tau^2}{\partial J_l}\right)t\int\frac{[2-\cos(2\pi x)]}{x^2}\rho(\nu)dx, \qquad (33)$$

where the upper limit in the integration is $+\infty$, and the lower limit is strictly speaking $-\omega_\tau t$, but if t is sufficiently large may in the usual manner be without sensible error treated as $-\infty$. We should in accordance with the definition of the symbol Σ' also consider in (33) terms which come from having $-\nu$ everywhere replaced by $+\nu$, but if now x denotes $(\nu+\omega_\tau)t$, these terms would lead to integrals of the same form as those given above except with the lower limit $+\omega_\tau t$. This means that if t is large the integration would be only over very large values of x, and

because of the factors x^{-2} or x^{-3} these integrals would be so small as to be negligible. This omission of terms in $\omega_\tau+\nu$ is, in fact, customary, and simply another way of saying that the only appreciable contributions to the perturbations come when the factors in the denominator approach zero, indicating resonance.

In considering the second integral of (33), $\rho(\nu)$ is a function of x, but may, following the usual procedure, be replaced by its value $\rho(\omega_\tau)$ at the resonance point $\nu=\omega_\tau$.[36] Since

$$\int_{-\infty}^{+\infty} \frac{1-\cos 2\pi x}{x^2}\,dx = 4\int_0^{\infty} (\sin^2\pi x)x^{-2}dx = 2\pi^2$$

the second integral in (33) becomes $2\pi^2\rho(\omega_\tau)$. A partial integration converts the first integral in (33) into one identical with the second[37] except that $\rho(\nu)$ is replaced by $\partial\rho/\partial x$, which is equal to $t^{-1}\partial\rho/\partial\nu$. The former integral is therefore $2\pi^2t^{-1}\partial\rho/\partial\nu$, where $\partial\rho/\partial\nu$ is to be evaluated at $\nu=\omega_\tau$.

Using the values just obtained for the integrals in (33), and summing over all the terms in the multiple Fourier expansion to get the total perturbation, we have

$$\overline{\Delta J_k}=\tfrac{2}{3}\pi^3e^2t\Sigma_\tau\left\{\tau_k D_\tau^2\left(\Sigma_{(l=1,2,3)}\tau_l\frac{\partial\omega_\tau}{\partial J_l}\right)\frac{\partial\rho}{\partial\nu}+\tau_k\left(\Sigma_{(l=1,2,3)}\tau_l\frac{\partial D_\tau^2}{\partial J_l}\right)\rho(\omega_\tau)\right\} \quad (34)$$

In writing this equation we have everywhere replaced X_τ^2 by the expression $D_\tau^2=X_\tau^2+Y_\tau^2+Z_\tau^2$ in order to allow for the perturbations produced by the y and z components of the electric intensity. It is obvious from symmetry considerations that taken by itself the y component of the electric intensity produces perturbations in which X_τ is everywhere replaced by Y_τ, with an analogous substitution for the z component. It is also clear that the total average perturbation is the sum of the x, y, and z effects taken separately, for cross terms involving factors such as E_xE_y cancel out on the average if the phase relations of the impressed radiation are haphazard. Performing this sum amounts simply to the substitution of D_τ^2 for X_τ^2.

In Eq. (34) it is to be understood that $\partial\rho/\partial\nu$ and $\rho(\nu)$ are both to be evaluated with the argument $\nu=\omega_\tau$. This means that these expressions have different values for each term in the summation. A similar evalua-

[36] Cf. Planck, Wärmestrahlung, 4th Ed., pp. 151 and 164. A Taylor's expansion of $\rho(\nu)$ in powers of $\nu-\omega_\tau$ shows that the departure of $\rho(\nu)$ from $\rho(\omega_\tau)$ may be disregarded in evaluating the 2nd (but not the 1st) integral of (33), as the terms thus neglected involve powers of t in the denominator.

[37] The integrand of the 1st integral in (33) may be written as $\rho\partial V/\partial x$ where $V=-(1-\cos x)/x^2$.

tion of $\partial\rho/\partial\nu$ and $\rho(\nu)$, varying with each term, is also to be understood in all subsequent τ summations.

12. Calculation of $\overline{\Delta J_k \Delta J_l}$

The next step is to compute the average value of products of the type $\Delta J_k \Delta J_l$, as a knowledge of this kind of expression to 2nd powers of E_x is necessary for a determination of the absorption and also for the analysis in part III. To make the calculation we multiply together the series for ΔJ_k and ΔJ_l given in Eq. (29).[38] For reasons similar to those explained in section 10, the only products which need be retained in taking the average are those which arise from the multiplication of a term in $\omega_\tau \mp \nu$ by one in $-\omega_\tau \pm \nu$. Following a procedure closely analogous to the derivation of (32) we thus find

$$\overline{\Delta J_k \Delta J_l} = \tfrac{1}{4} e^2 E_x^2 \, \Sigma'_\tau \left\{ \tau_k \tau_l X_\tau^2 \, \frac{1-\cos 2\pi(\omega_\tau - \nu)t}{(\omega_\tau - \nu)^2} \right\}. \tag{35}$$

The passage from a monochromatic plane wave to continuous homogeneous radiation is identical in character with the transition from (32) to (34). In fact the above expression is similar in the form of variation with ν to the second term of the right hand side of (32), and we can by inspection immediately write down the final result analogous to (34)

$$\overline{\Delta J_k \Delta J_l} = \tfrac{4}{3}\pi^3 e^2 t \Sigma_\tau \left\{ \tau_k \tau_l D_\tau^2 \rho(\omega_\tau) \right\}. \tag{36}$$

13. Calculation of Energy Absorbed

It is now easy to compute the absorbed energy, which we shall denote by $W - W_0$. The total energy (kinetic+potential) is identical with the Hamiltonian function defined in (20), and by postulate (III) concerning the J,w coordinate system, the latter can be expressed in terms of the J's alone, without involving the w's. Therefore

$$W - W_0 = H(J_1^0 + \Delta J_1,\ J_2^0 + \Delta J_2,\ J_3^0 + \Delta J_3) - H(J_1^0,\ J_2^0,\ J_3^0).$$

Expanding in a power series in the ΔJ's, and keeping terms of the order of magnitude of the square of the electric field, we have

$$W - W_0 = \sum_{(k=1,2,3)} (\overline{\omega_k \Delta J_k}) + \tfrac{1}{2} \sum_{(k=1,2,3)} \sum_{(l=1,2,3)} \left[\left(\frac{\partial \omega_k}{\partial J_l} \right) \overline{\Delta J_k \Delta J_l} \right] \tag{37}$$

where, as previously, we have utilized the fact that $\omega_k = (\partial H/\partial J_k)_0$, etc. Bars have been added to indicate that the average is to be taken over a

[38] Only the expressions for ΔJ_k and ΔJ_l accurate to first powers of E_x need be employed, as we desire their product to squares of E_x.

large number of waves. The values of ΔJ_k and $\Delta J_k \Delta J_l$ given in (34) and (36) may now be substituted. When this is done, and the summations with respect to l are expanded, the formula (37) becomes

$$W - W_0 = \tfrac{2}{3}\pi^3 e^2 l \Sigma_\tau \left\{ \rho(\omega_\tau) \left(\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \tau_3 \frac{\partial}{\partial J_3} \right) G_\tau + \frac{\partial \rho}{\partial \nu} G_\tau \left(\tau_1 \frac{\partial \omega_\tau}{\partial J_1} + \tau_2 \frac{\partial \omega_\tau}{\partial J_2} + \tau_3 \frac{\partial \omega_\tau}{\partial J_3} \right) \right\} \tag{38}$$

where we have written G_τ for $D_\tau \omega_\tau$. This is the same result as was quoted in Eq. (16) of part I.[39]

It is readily verified that the general result given in (38) reduces to the values for the absorption derived by Planck and Fokker respectively should the system be a linear oscillator or a rotating dipole.[40]

14. Study of the Special Case where $\partial\rho/\partial\nu = 0$

A particularly interesting situation arises when the radiation field is such that the specific energy $\rho(\nu)$ is independent of the frequency, making $\partial\rho/\partial\nu = 0$, and enabling one to take $\rho(\omega_\tau)$ outside of the summation sign in Eq. (38).

[39] In Eq. (16) the effect of only the resonance to one particular term in the multiple Fourier expansion was considered, and a factor of N was inserted to give the absorption by N atoms rather than only one. Also in (16) the time interval was denoted by Δt rather than t. These facts immediately explain superficial differences between the forms of (16) and (38).

[40] For an oscillator there is just one frequency, so that terms in τ_2 and τ_3 do not enter, and we can only have $\tau_1 = 1$ as there are no harmonics. Also $\partial\omega/\partial J$ is zero as the frequency is independent of the size of the orbit, while G_τ is equal to $J/2\pi^2 m$, as is readily proved by reasoning similar to that used in the derivation of (12), except that nh is replaced by J. Therefore $W - W_0 = \pi e^2 \rho(\nu) t/3m$, in agreement with Planck and others (cf. Eq. 14).

For a dipole rotating in a plane we may as above take $\tau_2 = \tau_3 = 0$, while τ_1 must equal 1. Also $J = 4\pi^2 I\omega$, where I is the moment of inertia. (This becomes obvious when we remember that the J which on the basis of the quantum theory is to be equated to nh is the angular momentum $2\pi I\omega$ multiplied by 2π.) If l denotes the length of the dipole, the motion is of the form $x = e^{-1} l \cos 2\pi\omega t$, $y = e^{-1} l \sin 2\pi\omega t$, so that $X_\tau = Y_\tau = l/e$, and $G_\tau = 2\omega l^2 e^{-2} = Jl^2/2\pi^2 e^2 I$. Therefore by (38) $W - W_0 = \frac{1}{3}\pi l^2 I^{-1}[\rho(\omega) + \omega(\partial\rho/\partial\nu)]$. This is the result obtained by Fokker (Ann. der Phys. **43**, 818). (Fokker does not actually write down the absorption, but the latter is simply computed from his expression for the fluctuation in the angular momentum. In a comparison with the present article, allowance must be made for the fact that Fokker uses Heaviside-Lorentz units of charge).

The preceding two examples illustrate how easily the results for special concrete cases can be derived from the general formula (38). This Eq. for the absorption might also have been computed as the time integral of $-E_x \dot{e} x \cos 2\pi(\nu t + \epsilon_0)$. This would be easier if we did not have to know the ΔJ's to second order terms for the work of part III.

It will be shown elsewhere by the writer[41] that $J_k = \partial \bar{T}/\partial \omega_k$, where $\bar{T}$ denotes the time average of the kinetic energy, and where in the differentiation the frequencies ω_1, ω_2, ω_3 are to be regarded as variables, but all the amplitudes X_τ, etc., are held fast. The formula just cited for J_k may be regarded as a method for defining the J's which is alternative and equivalent to Eq. (2). Now since $\bar{T} = m\pi^2 \Sigma_\tau \omega_\tau^2 D_\tau^2$ and $\omega_\tau = \tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3$, we have

$$J_k = \partial T/\partial \omega_k = 2m\pi^2 \Sigma_\tau \{ \tau_k G_\tau \} \tag{39}$$

where $G_\tau = \omega_\tau D_\tau^2$. Therefore when $\rho(\omega_\tau)$ is a constant independent of ω_τ, (38) takes the very simple form

$$W - W_0 = \pi e^2 \rho t/m \ . \tag{40}$$

Thus we see that when $\partial\rho/\partial\nu = 0$, the mean absorption of each electron is $\pi e^2 \rho t/m$, regardless of the shape and size of the orbit or the character of the dynamical system to which the electron belongs. That is to say, in a field of radiation whose specific energy does not vary with the frequency, we have the rather surprising result that the mean absorption is independent of the form of the force function $V(x, y, z)$ which holds the electron in the atom.[42, 43]

[41] The Fundamental Concepts of the Quantum Theory of Line Spectra, Appendix 1, to appear as a Bull. Nat. Research Council. Eq. (39) may also be derived from a somewhat analogous relation given by Burgers (viz., Eq. 9, p. 47 of his dissertation, "Het Atoom-model van Rutherford-Bohr").

[42] The result (40) also holds for free electrons subject to a field for which $\partial\rho/\partial\nu = 0$, regardless of whether or not free paths are stipulated. This can be seen from the analysis given by Jeans (Phil. Mag. **17**, 773, 1909) if his expressions for the absorption (his Eq. (34) and equation at bottom of p. 790) are integrated over all frequencies, regarding $\rho(\nu)$ as constant, and taking the dielectric constant as unity. In the case of free paths, the result may also be derived from an earlier paper by Sir J. J. Thompson (Phil. Mag. **14**, 223, 1907). If the orbital electron is constrained in such a way as to have only one or two degrees of freedom, we must clearly introduce factors of 1/3 or 2/3 into (40). The result thus obtained in the case of one degree of freedom is the same as the absorption by a linear oscillator in a variable field such that the density $\rho(\nu)$ at the resonance point is just equal to the constant ρ employed above (cf. Eq. 14).

[43] It is tempting (but probably futile) to determine the probability coefficient $A(r \to s)$ in (5) in such a way that a relation similar to (40) is preserved when the quantum theory is substituted for classical electrodynamics. The quantum theory analogue of the total absorption is the sum of all the different kinds of differential absorption (cf. sections 4 and 5) of various frequencies which may commence at a given orbit. If in Eq. (10) we use the + signs and take $m = 1$, it can be verified without particular difficulty with the aid of (39) and (7) that even at ordinary wave-lengths or small quantum numbers we have in the quantum theory a relation of the form (40), as the right hand side of Eq. (15) becomes $\pi e^2 \rho N \Delta t/m$ when summed over all values of r and t. (Such a relation must hold asymptotically for large quantum numbers regardless of how $A(r \to s)$ is evaluated, provided only the correspondence principle for emission is valid.) However, such a method of determining $A(r \to s)$ is hard to reconcile with the work of F. C. Hoyt (loc. cit.

15. Computation of Polarization

In order to furnish a basis for statements made in the discussion of dispersion given in section 6, it is necessary to calculate the polarization of the atom produced by the incident waves. It is clear from a power series expansion of Eq. (21), that to first order terms in the electric field the x component of the displacement of the electron brought about by the impressed waves is given by[44]

$$\Delta x = \Sigma_\tau \left\{ \sum_{(l=1,2,3)} \left(\frac{\partial A_\tau}{\partial J_l} \Delta J_l + 2\pi i \tau_l \Delta w_l A_\tau \right) \exp[2\pi i(\omega_\tau t + \epsilon_\tau)] \right\}.$$

The values of the ΔJ's and Δw's given by Eqs. (29) and (30) can now be substituted. The displacement x is thus expressed as the product of two multiple Fourier series. In multiplying these series together term by term we need only retain the products which involve the time just through a factor $\exp[2\pi i \nu t]$ (meaning that terms, for instance, in $\exp[2\pi i(\omega_\tau + \omega_{\tau'} + \nu)t]$ etc. may be disregarded). The reason for this is that in considering dispersion we are interested in the energy which is re-radiated by the atom as light of the same frequency as the impressed wave, and not in energy which is dissipated in light of the characteristic frequencies of the atom—a sort of fluorescence effect.[45] This means that the only products of significance for our purposes are those which arise from the multiplication of a term in ω_τ by one in $-\omega_\tau$, and furthermore in multiplying together such a pair of terms we may disregard the constant and secular parts[46] of ΔJ or Δw and retain only the portions of the right hand

---

in footnote 6), who showed that much better agreement with a limited amount of x-ray data is obtained with $m=4$ than with $m=0$. The case $m=1$ would probably resemble more closely $m=0$ than $m=4$. Another objection to trying to get an exact quantum theory analogue of (40) is that in so doing it is necessary to sum over all + and − integral values of $\Delta J_k/h$ ($k=1, 2, 3$) and some of these would in general lead to transitions from positive to negative quantum numbers, which can scarcely correspond to any physical reality. Also it is doubtful how much real physical significance can be attached to radiation fields such that $\partial\rho/\partial\nu=0$.

[44] Higher order terms in the power series expansion would give a change in the dielectric constant with field strength which may be disregarded for weak fields.

[45] Combinations $\omega_\tau \pm \nu$ of the natural and impressed frequencies also appear in the term by term multiplication and would be radiated. However these terms are not of direct interest in dispersion, as their frequency differs from ν. Frequencies of the form ω_τ are of course radiated even in the absence of an impressed field, and the total amount of energy emitted is only slightly modified by the addition of the latter.

[46] By "secular" terms are meant those which involve the time explicitly rather than through exponentials. Such terms are found in (30) but not (29). The effect of these terms in (30) is simply to change slightly the vibration frequencies of the system, rather than to introduce new vibrations resonant to the impressed wave.

sides of (29) and (30) which involve the time through exponentials. Remembering that the prime in the symbol Σ' signifies that the summation is to be repeated with $-\nu$ and $-\epsilon_0$ everywhere replaced by $+\nu$ and $+\epsilon_0$, we thus get for the part of the displacement which is resonant to the impressed wave[47]

$$\Delta x = \tfrac{1}{2} e E_x \Sigma_\tau' \left\{ \sum_{(l=1,2,3)} \left[-\tau_l \left(\frac{\partial A_\tau}{\partial J_l} A_{-\tau} + A_\tau \frac{\partial A_{-\tau}}{\partial J_l} \right) + \frac{\partial \omega_\tau}{\partial J_l} A_\tau A_{-\tau} (\omega_\tau - \nu)^{-1} \right] \frac{\exp[2\pi i(\nu t + \epsilon_0)]}{\omega_\tau - \nu} \right\}.$$

By combining terms in $\omega_\tau \mp \nu$ with those in $-\omega_\tau \pm \nu$ and using (25) we obtain by a procedure similar to that used in deriving (32)

$$\Delta x = \tfrac{1}{4}\, e\, \mathscr{E} \sum_\tau{}' \sum_{(l=1,2,3)} \left[-\tau_l \frac{\partial X_\tau^2}{\partial J_l} (\omega_\tau - \nu)^{-1} + \frac{\partial \omega_\tau}{\partial J_l} X_\tau^2 (\omega_\tau - \nu)^{-2} \right]$$

where now the number of terms in the τ summation is halved by the requirement that ω_τ be positive, and where $\mathscr{E}$ has been written for $E_x \cos 2\pi(\nu t + \epsilon_0)$.

To get the polarization we multiply Δx by the electronic charge $-e$ and the number N_s of atoms per unit volume. On combining terms in $\omega_\tau - \nu$ and $\omega_\tau + \nu$, and making some slight rearrangements, this gives

$$P_x = -eN_s \Delta x = N_s e^2 \mathscr{E} \sum_\tau \sum_{(l=1,2,3)} \left[\frac{\tau_l \frac{\partial}{\partial J_l} (\omega_\tau X_\tau^2)}{2(\omega_\tau^2 - \nu^2)} - \frac{\omega_\tau^2 X_\tau^2 \tau_l (\partial \omega_\tau / \partial J_l)}{(\omega_\tau^2 - \nu^2)^2} \right]. \tag{41}$$

Besides P_x, which is the x component of the polarization due to the electric intensity $\mathscr{E}$, there would also be y and z components P_y and P_z, involving products of the form $A_{-\tau}B_\tau$, $A_{-\tau}C_\tau$, etc. where B_τ and C_τ are coefficients analogous to A_τ except that they appear respectively in the expansions of the y and z components of the unperturbed motion, rather than that for x. This would give an anisotropic effect similar to that found in a crystal, as $\mathscr{E}$ has by hypothesis the x direction, while P would be differently directed. However, this is on the assumption that all N_s atoms are entirely similar,—i.e. have internal force fields with the same spacial orientations. If we assume that the axes fixing the force function $V(x, y, z)$ have random distribution,[48] then on the average $\overline{X_\tau^2} = \tfrac{1}{3} D_\tau^2$, while $\overline{A_{-\tau}B_\tau} = \overline{A_{-\tau}C_\tau} = 0$, making $P_y = P_z = 0$ and eliminating the anisotropy.

[47] In forming a term in $-\omega_\tau$ in (29) or (30) care must be taken to reverse the sign of τ_l or τ_k as well as just ω_τ.

[48] Cf. footnote[45]

On making the substitution $\frac{1}{3}G_\tau=\frac{1}{3}\omega_\tau D_\tau^2$ for $\omega_\tau X_\tau^2$, expanding the l summation, and taking only one term in the τ summation so as to give just the resonance between the impressed frequency and one particular combination overtone ω_τ, (41) becomes identical with Eq. (18) of part I. This is the desired result.[49]

By reasoning similar to that given in footnote[40] it may easily be verified that for linear oscillators whose axes are parallel to the electric intensity, (41) reduces to the well-known result $(N_s e^2\mathscr{E}/4\pi^2 m)\,(\omega^2-\nu^2)^{-1}$. For dipoles whose axes of rotation have random spacial orientations, the polarization is $-(N_s l^2\mathscr{E}/12\pi^2 I)\,(\nu^2+\omega^2)\,(\omega^2-\nu^2)^{-2}$ (notation as in footnote 40).[50]

16. Calculation of Distortion of Orbit by Forces due to Electron's Own Radiation

In the preceding sections we have calculated the perturbations in the orbits produced by impressed radiation fields. It is well known that there is also a force which is due to the electron's own radiation. Neglecting higher powers of $1/c^2$, this force has[51] the value $2e^2\ddot{v}/3c^3$, where v denotes the vector velocity. From Eqs. (21) and (24) it follows that the x component is

$$F_x=(-16\pi^3 ie^2/3c^3)\,\Sigma_\tau\{\omega_\tau^3 A_\tau\exp[2\pi i(\omega_\tau t+\epsilon_\tau)]\}\ . \tag{42}$$

Strictly speaking the force will not have exactly this value, for in differentiating the expression for x three times to get $\ddot{v}$, we have neglected the slow alterations in A_τ and ω_τ produced by the damping due to the

[49] The calculation which we have given of the polarization is the simplest possible one. A refinement which should be considered in a more extensive analysis is the effect of the damping due to the electron's own radiation on the form of the dispersion curve near the resonance point. Also the modifications appropriate to degenerate systems would be of importance, as atomic systems are of this type in the absence of constant external fields. (We have avoided the degeneracy difficulty by introducing the artifice of internal force fields of random orientations). Considerations of such details are, however, beyond the scope of the present paper.

[50] It is to be clearly understood that in the computation of the polarization (as well as the rest of part II) it is assumed all orbits have the same size and shape. This would not be the equilibrium distribution, as it is shown in part III that this is given by the Boltzmann formula, leading to a continuous distribution of sizes and shapes. The latter, however, would give rise to continuous black body spectra rather than monochromatic phenomena, and would not explain the occurence of resonance frequencies in dispersion. Therefore in comparing the classical and quantum theories to establish correspondence principles, we must assume a discrete distribution even in the classical theory. An initial discrete distribution would according to the classical theory eventually become continuous due to the action of black body radiation and spontaneous emission. The present computations apply to a time interval long compared to a period but short enough so that the departures from the initial discrete distribution are not sensible.

[51] Cf. M. Abraham, Theorie der Elektrizität (4th Ed.), Vol. II, Eq. (87).

electron's own radiation. However introducing these corrections would clearly be a second order effect which can be neglected if the time interval is not too long.[52]

It is clear that the mathematical analysis for the present case resembles that for an impressed monochromatic wave, already studied in sections 8 and 9, except that now the force instead of being due to just one wave $-eE_x\cos[2\pi(\nu t+\epsilon_0)]$ is due to the infinity of waves of different frequencies given by (42).[53] Therefore in Eq. (27) or (28) for determining the perturbations we must replace $-\frac{1}{2}eE_x\{\exp[2\pi i(\nu t+\epsilon_0)]+\exp[-2\pi(\nu t+\epsilon_0)]\}$ by the expression F_x given above. We thus have in place of (28)

$$dJ_k/dt=2\pi iF_x\Sigma_\tau\{\tau_k A_\tau\exp[2\pi i(\omega_\tau t+\epsilon_\tau)]\}\ .$$

When the explicit value for F_x is substituted and the series are multiplied together, the "cross-terms" vanish on the time average, and the only terms which remain come from the multiplication of a term in ω_τ by one in $-\omega_\tau$. Consequently the time average of dJ_k/dt is $(-32\pi^4e^2/3c^3)\times\Sigma_\tau\tau_k\omega_\tau^3A_\tau A_{-\tau}$. If we utilize (25) and halve the number of terms in the τ summation by taking only such values of the integers τ_1, τ_2, τ_3 as make the expression ω_τ positive[54] we have

$$\overline{dJ_k/dt}=(-16\pi^4e^2/3c^3)\Sigma_\tau\{\tau_k\omega_\tau^3X_\tau^2\}\ . \tag{43}$$

Integration of this relation gives

$$\Delta J_k=t(-16\pi^4e^2/3c^3)\Sigma_\tau\{\tau_k\omega_\tau^3D_\tau^2\}\ . \tag{44}$$

In this equation we have inserted D_τ for X_τ in order to take into account also the y and z components of the radiation force. It is clear from symmetry considerations that the effect of these components is simply to add terms in Y_τ^2 and Z_τ^2. (Cross terms involving products such as $X_\tau Y_\tau$ cannot arise in the first order approximation we are seeking, as the Hamiltonian function is $H-xF_x-yF_y-zF_z$, where H is defined as in (20); and consequently the x, y, z effects enter separately unless squares of the perturbations are considered).

From (44) it follows that if we considered only the radiation force arising from one particular term ω_τ in the multiple Fourier expansion of F_x, the ΔJ's would be in the ratio

$$\Delta J_1:\Delta J_2:\Delta J_3=\tau_1:\tau_2:\tau_3\ . \tag{45}$$

[52] The interval must in a later part of the proof at the same time be assumed long compared with the natural periods of motion, but the addition of this restriction involves no contradiction. Cf. M. Planck, "Wärmestrahlung," p. 154 (4th Ed.).

[53] The transition from an equation of type (27) to one similar to (28) may be made just as in the monochromatic case, as the coordinates w_k and the amplitudes may to the order of approximation under consideration be replaced by their unperturbed values $\omega_k t+\epsilon_k$ and A_τ, just as was done in section 9.

[54] Cf. derivation of Eq. (32).

This result has already been quoted in Eq. (11) of part I.[55]

The average rate of radiation is the negative of the rate of change of the system's energy and is therefore on the average (cf. section 13)

$$-dW/dt = -\sum_{(l=1,2,3)} \frac{\partial H}{\partial J_l}\frac{dJ_l}{dt} = (16\pi^4 e^2/3c^3)\ \Sigma_\tau\ \omega_\tau^4 D_\tau^2 \tag{46}$$

as $\partial H/\partial J_k = \omega_k$ and $\omega_\tau = \tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3$ This is a well known result already utilized in Eq. (8) of part I. If only the combination overtone ω_τ were operative in the radiation force given in (42) only the corresponding term would appear in the expression given above for the rate of radiation. Viewed from this standpoint, the result stated in (45) seems almost intuitive; for since $dW/dt = \Sigma_{(l=1,2,3)}\ \omega_l dJ_l/dt$, it is hard to see how the formula for the radiation of energy could be of the well known type given in the right of (46) unless the J's were altered in the manner given in (43) (with X_τ replaced by D_τ) and hence in the ratios given in (45), when only one term is considered.

17. Generalization to Case of More than Three Degrees of Freedom

So far the results have for simplicity been stated only for the case of three J's and hence three intrinsic frequencies $\omega_1, \omega_2, \omega_3$. The case of more than three of these expressions can arise only when there are more than three degrees of freedom and consequently more than one electron or moving particle in the system. It is found without difficulty that in this case the Hamiltonian function in the presence of a monochromatic wave vibrating parallel to the x axis is no longer $H + xeE_x \cos[2\pi(\nu t + \epsilon_0)]$ (cf. Eq. 25b) but instead $H - \Sigma_j x_j e_j E_x \cos[2\pi(\nu t + \epsilon_0)]$ where the summation is to be extended over all the particles in the system, and x_j and e_j denote the coordinate and charge of the jth particle. The analysis can be carried through just as for one particle except that $-eA_\tau, -eX_\tau$, etc.,

[55] It is to be noted that the result (45) is not restricted to the case of the particular type of force obtained by taking one pair of conjugate terms in (42) but is easily proved equally valid for any other force of the type form $S\cos 2\pi\omega_\tau t + T\sin 2\pi\omega_\tau t$ where S and T are arbitrary constant vectors. This immediately indicates the possibility of proving (45) even when higher order terms in $1/c^2$ and the changes in ω_τ and A_τ with time are considered, all of which were neglected in deriving (42). For when these effects are considered and we keep just one combination overtone (terms in ω_τ and $-\omega_\tau$) in the resulting expansion for F_x, more rigorous than (42), we still have $\overline{dJ}_1/dt : \overline{dJ}_2/dt : \overline{dJ}_3/dt = \tau_1 : \tau_2 : \tau_3$. Since the J's always change instantaneously in this ratio (neglecting small periodic oscillations), we have (45) valid for an arbitrary time interval, even though (44) is an approximation valid only for intervals not long enough for the amount of damping to be appreciable.

must be replaced by the coefficients in the multiple Fourier expansion of $\Sigma_j e_j x_j$, which is the x component of the electrical moment of the atom. The results of sections 11, 12, 13, 15 are therefore unaltered except that, of course, allowance must be made for the fact that there are μ rather than three degrees of freedom by adding terms in $J_4, J_5, \ldots J_\mu, \tau_4, \tau_5, \ldots \mu$ and in the final equations $-eD_\tau$ must signify an amplitude in the multiple Fourier expansion of the polarization of the entire atom rather than the displacement of just one charge.[56]

It is also verified without difficulty that the content of section 16 and the correspondence theorems stated in sections 2, 5, 6 of part II can be extended to include the general case if only the modifications indicated in the preceding paragraph are introduced.[56]

DEPARTMENT OF PHYSICS
UNIVERSITY OF MINNESOTA,
June 19, 1924.

[56] All of the generalizations given above are on the assumption that the dimensions of the atom are small compared to the wave-length of the impressed beam. This situation has, in fact, been tacitly assumed throughout part II (even when there is only one electron), as it is required for the validity of the fundamental Eq. (25a). The case where the impressed wave-length is comparable with the dimensions of the orbits could be treated after the general methods of the present article, but the calculations would be probably rather laborious.

In atoms with several electrons a case which is readily treated is that where each orbit is itself small compared to the impressed wave-length but the distance between orbits is large compared to the latter. The electrons then absorb independently whereas the generalizations given above are essentially for the case where they absorb in unison. The appropriate modification consists in introducing $\Sigma_j e_j^2 D_j^2$ rather than $(\Sigma_j e_j D_j)^2$ in place of e^2D^2. This case of independent absorption could not well actually arise, as atomic dimensions are usually too small (except for x-rays) and also electrons travel all over the atom rather than being confined to small regions or cells. In a system with several electrons the absorption per electron in a constant radiation field is $\pi e^2 \rho t/m$ (this statement is the generalization of section 14) only if the electrons absorb independently rather than in unison.

It is interesting to note that when the wave-length radiated is large compared to atomic dimensions, the radiation force $2e^2\ddot{v}/3c^3$ affects not only the electron for which $\ddot{v}$ is computed, but also all other electrons in the atom. For if each electron were affected by its own radiation force but not that of the other electrons, the work done by the forces would be proportional to $\Sigma e_j^2 D_j^2$ rather than $(\Sigma e_j D_j)^2$, whereas it is well known that the loss of energy should involve the amplitudes in the latter rather than in the former fashion if the atom is small compared to the wave-length.

The assumption that the wave-length is large compared to the orbit cannot be regarded as valid when we consider the "higher harmonics" or combination overtones of extremely large frequency. These terms, however, will not contribute appreciably to the total absorption or radiation if the series converges sufficiently rapidly.

Because of the substitution of $\Sigma_j (e_j D_j)^2$ or $\Sigma_j e_j^2 D_j^2$ for e^2D^2 in (9) or (10) it should be noted that in atoms with more than one electron the probability coefficient $A_{\tau \to \sigma}$ depends in general on the amplitudes of all the different electrons rather than just one of them.

THE VARIATION OF THERMIONIC EMISSION WITH TEMPERATURE AND THE CONCENTRATION OF FREE ELECTRONS WITHIN CONDUCTORS*

By A. T. Waterman

Abstract

Equilibrium electron theory of thermionic emission.—The equation for the thermionic saturation current per unit area in the form $I = AT^{\alpha} e^{-b/T}$ is derived according to Richardson's thermodynamical method, assuming (1) the boundary condition that when ψ the free electron evaporation constant is zero, the free electron concentration inside and outside the hot body is the same, and (2) the interior free electron concentration given by the author's equilibrium theory of conduction (Phys. Rev. **22**, 259, 1923). α depends only upon the valence in the electron-forming reaction and always lies between 1/2 and 5/4. β depends upon the valence and the two evaporation constants ϕ_0 and ψ_0, referring respectively to a bound and a free electron. A is found to depend upon the valence, the electronic chemical constant and the concentration of electron-forming particles in the hot body, varying from $4.16(10)^{14}$ for BaO to $8.13(10)^{14}$ for Ni. Values of β calculated from the computed A and experimental data show excellent agreement with β observed for Pt, W, Mo, Ni, Ca, CaO, SrO, and BaO, the verification being more general than by use of Dushman's formula. The experimental data however are not sufficiently accurate to decide between the two formulas. The failure of certain hot bodies to agree with either formula is discussed and possible explanations offered. Assuming the theory correct, the *free electron concentrations in several metals, MoS_2 and CaO are calculated from experimental data.* For metals at 0°C this concentration is of the order of magnitude of 10^{17} per cc, and decreases with rising temperature. A noteworthy consequence of this low concentration is that the contribution of the free electrons to the specific heat of the conductor remains negligible down to very low temperatures.

THE variation of the thermionic saturation current density with temperature is given empirically by an expression of the form

$$I = AT^{\lambda} e^{-b/T} \tag{1}$$

where the exact value of λ matters little in obtaining agreement with experiment, provided it is of the order of magnitude of unity. The well-known theoretical expressions are[1]

$$I = A_1 T^{1/2} e^{-\psi/kT} \tag{2}$$

* The main conclusions of this paper were presented before the American Physical Society on Dec. 27, 1923 (Phys. Rev. **23**, 299, 1924)

[1] Richardson, Emission of Electricity from Hot Bodies, p. 36. ψ is used to denote the thermionic work function commonly denoted by ϕ, since in this paper ϕ is reserved for the photo-electric work function.

and $$I = A_2 T^2 e^{-\psi_0/kT} \tag{3}$$

derived respectively from kinetic theory and from thermodynamical considerations. The temperature functions in the two expressions become identical when the thermionic work function ψ is assumed to depend upon temperature according to the equation[2] $\psi = \psi_0 + (3/2)kT$, also derived from thermodynamical reasoning. The same form of equation has been obtained by application of the quantum theory.[3] Although the experimental evidence has thus far not been able to distinguish between Eqs. (2) and (3), Richardson and others have agreed that probably Eq. (3) deserves more confidence from the theoretical standpoint.

The evaluation of the constant A is a matter of interest and importance. Differences in the form of this constant arise chiefly from the different methods by which the concentration of external electrons in equilibrium with the hot body is deduced. Thus we have

$$n_A = n_B e^{-\psi/kT} \tag{4}$$

and

$$n_A = CT^{3/2} e^{-\psi_0/kT} \tag{5}$$

corresponding respectively to Eqs. (2) and (3), where n_A = equilibrium concentration of external electrons, n_B = electron concentration within the hot body, and C involves a constant of integration.

Dushman,[4] by application of the Nernst heat theorem, has recently derived an expression for the thermionic saturation current density of the form of Eq. (3), by evaluating the constant C in Eq. (5), with the interesting result that C, and hence A_2, should be universal constants. However the premises on which this latter derivation is based may possibly be open to question, and furthermore the universality of A_2 does not appear to extend to all possible emitting substances. The author has submitted[5] an expression for the thermionic saturation current density which is also consistent with Eq. (1). The previously omitted details of derivation are presented below, and the work has been extended to include evaluation of the integration constant appearing in the coefficient A (Eq. 1). This makes possible a numerical calculation of the value of A from data independent of thermionics, and thus permits a more substantial examination as to the validity of the formula proposed, as

[2] Richardson, loc. cit,[1] p. 35
[3] Richardson, loc. cit.[1] p. 41.
[4] Dushman, Phys. Rev. **21**, 623 (1923)
[5] Waterman, Phys. Rev. **22**, 259, 268 (1923)

well as allowing direct comparison with values of A derived by other methods.

Derivation of Expression for Thermionic Saturation Current Density

From the thermodynamic reasoning[6]

$$dn/n_A = \left(\frac{\psi}{kT^2}\right)dT \tag{6}$$

where ψ is the energy required to liberate one free electron from the hot body;[6] k is the gas constant per electron. Hence $\log n_A = \int(\psi/kT^2)dT +$ const. Using the boundary condition that, when $\psi = 0$, the concentration of free electrons inside and outside the hot body is the same, or $n_A = n_B$, we have

$$\log\frac{n_A}{n_B} = \int\frac{\psi}{kT^2}dT\,. \tag{7}$$

Since [2] $\psi = \psi_0 + (3/2)kT$

$$n_A = n_B T^{3/2} e^{-\psi_0/kT}\,. \tag{8}$$

From the writer's equilibrium theory of electrical conduction[7] the internal equilibrium concentration of free electrons n_B is given by

$$n_B = BT^{-3\nu/2(\nu+1)}\, e^{-\nu(\phi_0-\psi_0)/(\nu+1)kT} \tag{9}$$

where ϕ_0 and ψ_0 are respectively the energies (at 0°K) required to liberate a bound and a free electron from the conductor, whence $\phi_0 - \psi_0$ is seen to be the energy (at 0°K) required to change a bound electron into a free electron within the substance;

ν = valence in atomic reaction: Atom $\leftrightarrows$ positive ion $+ \nu$ electrons;

$B = c\,(\nu N)^{1/(\nu+1)}$;

N = concentration of nuclei within conductor (or total concentration of particles in conductor from which electrons may be or have been emitted).

The constant of integration c may be evaluated by noting that the integration constant of Eq. (3) in the writer's article[5] is for the general atomic reaction equal to νi, where i is the chemical constant of the electron, reckoned in absolute units per electron and in terms of concentrations instead of partial pressures.

Then c becomes $e^{\nu i/(\nu+1)}$, whence $B = e^{\nu i/(\nu+1)}(\nu N)^{1/(\nu+1)}$

[6] Richardson, loc. cit.[1] p. 31, Eq. (6)
[7] Waterman, loc. cit.[5] p. 259

Substituting the value of n_B from Eq. (9) in Eq. (8)

$$n_A = e^{\frac{\nu i}{\nu+1}} (\nu N)^{\frac{1}{\nu+1}} T^{\frac{3}{2(\nu+1)}} e^{-\frac{\nu\phi_0+\psi_0}{(\nu+1)kT}}, \tag{10}$$

giving the concentration of external electrons in equilibrium with the hot body. This concentration is seen to depend upon the concentration of nuclei or electron-forming particles within the substance, the valence in the atomic reaction, and the chemical constant of the electron, in addition to the temperature and the evaporation constants.

Following the customary procedure, the thermionic saturation current density now becomes

$$I = A T^{\alpha} e^{-\beta/T} \tag{11}$$

where

$A = \epsilon \sqrt{(k/2\pi m)}\ e^{\nu i/(\nu+1)} (\nu N)^{1/(\nu+1)}$ (ϵ being the electronic charge);

$\alpha = (\nu+4)/2(\nu+1)$;

$\beta = (\nu\phi_0+\psi_0)/(\nu+1)k$.

For different valences $\nu = 1, 2, 3$, the values of α are respectively $\alpha = 1.25$, 1.00, 0.875, and thus the exponent of T for all valences is seen to lie between the usual 1/2 and 2.

It should be observed that the constant β in Eq. (11) as experimentally determined does not necessarily yield directly the energy required to liberate a free electron from the substance (ϕ_0 as usually understood), as is true for b in the ordinary expression for the thermionic saturation current.[5] Instead β involves both the energy required to liberate a free electron and the energy required to liberate a bound electron. It is clear that such must be the case if the external concentration of electrons depends upon the internal concentration of *free* electrons and if the latter in turn is determined by a chemical equilibrium between atoms, positive ions and free electrons. Thus an amount of energy $(\phi_0-\psi_0)$ is required to change a bound electron into an internal free electron, and an amount of energy ψ_0 is then required to liberate this electron from the substance.

Evaluation of Constants and Application to Experimental Data

Using values of constants as follows:

$\epsilon = 4.774 \times (10)^{-10}$ e.s.u.; m (electron) $= 8.995 \times (10)^{-28}$gm.

$k = 1.372 \times (10)^{-16}$ ergs/deg. C; $i = 34.98$;

we have $\log A = -4.129 + 15.19\ \nu/(\nu+1) + [1/(\nu+1)] \log \nu N$.

In view of the uncertainty attaching to the theoretical values of the chemical constant of the electron, the value of the chemical constant

used is that selected by Tolman[8] from experimental data, expressed in terms of concentrations instead of partial pressures and in absolute units per electron. N, the concentration of nuclei within the conductor, may be determined by (density/atomic weight)×Avogadro's number.[9] The valence ν is the only arbitrary constant appearing in the expression, and for most substances agreement with experiment is obtained for $\nu=1$ (i.e. a univalent reaction liberating electrons within the conductor). Thus the value of the constant A may be calculated, after which an experimental value for the thermionic saturation current density at any single temperature enables β to be calculated. This computed value of β may then be compared with the value of β obtained directly from the slope of the graph ($\log I=\alpha \log T$) vs. $1/T$. The following table shows the results of this procedure for a number of emitting substances.

TABLE I

Hot body	Observer	Valence	α	A(calc.)	β(calc.)	β(obs.)	Dushman β(calc.)	Dushman β(obs.)	ψ_0
W	Davisson and Germer	1	1.25	$7.16(10)^{14}$	57,300* 56,800†	53,600* 54,600†	51,950*	52,000*	(volts)
		1.25	1.17	$2.98(10)^{14}$	54,100*	$54,000_1$*			4.66
	K. K. Smith	1	1.25	$7.16(10)^{14}$	52,150	53,700	53,900	51,400	
Ta	Deininger	3.3	0.85	$4.71(10)^{12}$	47,000	47,000	55,000	46,700	
Ca	Horton	2	1.00	$3.57(10)^{13}$	41,100	41,000	43,800	38,800	3.54
Mo	Stoeckle	1	1.25	$7.35(10)^{14}$	52,800	52,300	47,850	51,200	4.51
Pt	Schlichter	1	1.25	$7.28(10)^{14}$	53,100	52,500	49,150	50,000	4.53
	Deininger	1	"	"	51,100	50,000			4.31
Ni	Schlichter	1	1.25	$8.13(10)^{14}$	51,600	51,500	47,900	49,000	4.44
CaO	Jentsch	1	1.25	$5.33(10)^{14}$	40,000	39,500	37,000	38,900	1.6
	Deininger	"	"	"	40,750	42,700			($\phi_0=5.3$)
BaO	Jentsch	1	1.25	$4.16(10)^{14}$	41,000	41,000	38,000	39,700	
SrO	"	1	1.25	$4.24(10)^{14}$	44,250	44,200	41,100	42,200	

* Worthing and Forsyth temperature scale;
† Langmuir temperature scale.

The value given for β(calc.) in the case of any substance is the mean of separate calculations of β for each temperature at which emission was measured. Naturally, where agreement is close between β(calc.) and β(obs.), these separate values of β(calc.) showed a constancy within the limits of experimental error. Where a discrepancy occurs it is evident that the individual values of β(calc.) showed a progressive variation throughout the temperature range.

Where the accuracy of the experimental data seemed to warrant it, the current density was corrected for the expansion of the emitting surface

[8] Tolman, J. Am. Chem. Soc. **43**, 1592 (1921)
[9] In the case of the oxides the concentration of mole cules was used.

with temperature, since this may cause a difference in β (calc.) of 1 to 2 per cent.

It is seen from the table that a valence $\nu=1$ gives very satisfactory agreement for all hot bodies listed except for W, Ca and Ta. For W exact agreement is secured with $\nu=1.25$, i.e., some of the atoms may ionize doubly, giving a valence between 1 and 2. Ca requires a valence of 2, which appears reasonable, and for Ta ν apparently lies in the neighborhood of 3. The experimental data for Ta given by Deininger are not satisfactory however, in that the graph ($\log T-\alpha \log T$) vs. $1/T$ is not a straight line for any value of α. It may furthermore be noted that the values of A which give excellent agreement between β (calc.) and β (obs.), vary from $8.13\times(10)^{14}$ for Ni to $4.16\times(10)^{14}$ for BaO, even though the valence is the same (unity). However, as far as the effect of this variation in A on the value of β (calc.) is concerned, A might be considered approximately a universal constant for those cases where $\nu=1$, since taking the mean of A (calc.) for such cases would change the calculated value of β by only ± 2 per cent at the most. The obvious reason on this hypothesis for this approximate constancy of A for the same valence lies in the fact that the concentrations of nuclei in different elements and of molecules in compounds are closely of the same order of magnitude. However the use of the valence factor makes it possible to extend agreement to other substances, in which case a large change in A occurs.

Comparison with Other Formulas

For comparison the values of β (calc. and obs.) are included as computed according to Dushman's expression,[4] in which $\alpha=2$, and A is the universal constant $1.80\times(10)^{11}$ in the units employed in this table (e.s.u./cm (deg.)^2).

From the results contained in Table I it appears that the expression here proposed for the thermionic saturation current density compares favorably with that of Dushman, and in fact with any expression where A is taken to be a universal constant. It should be noted that Dushman's value of A for best agreement is based upon the chemical constant of the electron computed from the Sackur-Tetrode expression. The values of A computed in this paper however are based upon Lewis' and Gibson's calculation from the experimental determination of the entropy of helium, following Tolman.[8] The use of the chemical constant derived from Lewis-Gibson-Latimer theory[10] or the one selected by Tolman, when

[10] Lewis, Gibson, and Latimer, J. Am. Chem. Soc. **44**, 1008 (1922)

used for calculation of A in Dushman's expression, brings increasing disagreement between β (calc.) and β (obs.). As a matter of fact, the value of A in Dushman's equation which gives the best general agreement for all substances in Table I (omitting Ta) is about $5\times(10)^{11}$ instead of $1.8\times(10)^{11}$. The use of the other two values for the chemical constant results in values of A of $1.53\times(10)^{11}$ and $1.15\times(10)^{11}$ respectively. In the same units the value of A calculated from Richardson's application of the quantum theory is $1.5\times(10)^{10}$.

In justice to Dushman's equation it should be stated that, of the various substances listed, it shows its best agreement for the most reliable data. The expression derived in this paper on the other hand shows substantially the same agreement in these cases, but has a more extensive application.

From the theoretical point of view, aside from the evidence mentioned regarding the choice of chemical constant, the proposed formula for the saturation current seems to rest upon a more satisfactory physical basis, if we leave out of account the particular expression used for the internal electron concentration. On the basic assumption employed in derivation, quantitative results cannot be obtained unless some expression for this internal concentration is assumed. This very fact however may be advantageous in furnishing further evidence concerning the order of magnitude and variation with temperature of the electron concentration in conductors. Whether the expression here used for this internal concentration of free electrons, based upon a chemical equilibrium relation, is justified or not, can only be decided by the success of its application to this and to other electrical phenomena. The measure of its success in explaining the general temperature variation of electrical conductivity for both good and poor conductors has already been discussed.[5]

On the other hand the assumptions underlying Dushman's thermodynamical derivation may be questioned. Thus the hypothesis that thermionic emission is completely analogous to evaporation from a solid or liquid is in all probability merely an approximation. As the electrons are held in the solid only by virtue of the solid atoms, the case would seem to be more strictly analogous to the emission of absorbed gas from a solid or of dissolved gas from a liquid. In the latter case the equilibrium pressure of the external gas is directly proportional to the concentration of gas dissolved in the liquid. This condition is satisfied for the analogous case of electron emission in the present derivation whereas according to Dushman's conclusions the equilibrium pressure of an electron gas should be independent of the electron concentration within the emitting body.

Furthermore, in Dushman's derivation, the specific heat of the internal electrons c_p is assumed to be negligible as compared with that of the external electrons. However, the most probable reason why the internal electrons do not contribute appreciably to the specific heat of a metal, if there is any truth in the application of the kinetic theory to the internal electrons,[11] is that this concentration is small compared with the concentration of atoms. It is not beyond the bounds of possibility that the inclusion of the term involving c_p might produce an appreciable correction to the thermionic current as deduced by Dushman, even though the contribution of c_p to the specific heat of the hot body appears to be negligible. In this way the constant A of the saturation current, Eq. (3), might be found to vary slightly for different substances, and thus agreement be obtained between the two expressions Eqs. (3) and (11).

Values of the Work Functions ϕ_0 and ψ_0

In the last column of Table I are given various values of the energy functions ϕ_0 and ψ_0 which appear in β. As stated above ψ_0 represents the energy required (at 0°K) for emission of one *free* electron from the substance, while ϕ_0 is the energy necessary for emission of one *bound* electron. ϕ_0 has been identified[5] with the photo-electric work function, and ψ_0 with the work function entering into the usual theory of thermionic emission. In the case of metals calculation of ϕ_0 and ψ_0 from β may readily be made, since[5] the conductivity evidence indicates that ϕ_0 and ψ_0 do not differ by more than 1/10 percent. Thus $\beta=(\nu\phi_0+\psi_0)/(\nu+1)k=\psi_0/k$ (or ϕ_0/k) for these substances. But for poor conductors like the oxides ϕ_0 is substantially greater than ψ_0.[5] However, if from data on conduction,

[11] Note added July 19: It has been shown by Davisson and Germer (Phys. Rev. **20**, 329, 1922) and recently emphasized by Richardson (Proc. Roy. Soc. **105**, 387, 1924) that in the accurate work of the former on pure tungsten the values of the thermionic work function ψ (in the present notation) as obtained simultaneously by calorimetric and by thermionic methods agree only when the thermal energy of the internal electrons involved in thermionic emission is negligible. This is naturally accepted as important evidence against the classical free electron theory. It should be pointed out, however, that the argument as presented may fail if the constant A of the thermionic equation is a function of the temperature, which presumably is the case if the concentration of internal free electrons depends upon temperature. In particular, using the function of temperature employed in this paper for the internal free electron concentration the above conclusion is exactly reversed. Thus $\psi=\phi_0=4.66$ v (cf. Table I) for W, whence $\psi=\psi_0+(3/2)kT=4.66+0.29=4.95$ v at 2270°K. On attributing to the internal electrons the classical value for their thermal energy this number should agree with Davisson and Germer's calorimetric value of ψ, without correction, namely 4.91 v. The agreement is seen to be to less than 1 percent on the classical theory, while according to the non-classical point of view, ψ(thermionic) = 4.66 v and ψ(calorimetric) = 4.52 v, a discrepancy of 3 per cent.

$b = \nu(\phi_0 - \psi_0)/(\nu+1)k$ of the expression for specific resistance is found,[12] then the independent values of ϕ_0 and ψ_0 may be computed. For CaO, Horton's data[13] give $b = 21{,}000$ approximately, which taken in conjunction with $\beta = 40{,}000$ results in the values of ϕ_0 and ψ_0 given in the Table.

Discussion of Lack of Agreement for Most Oxides

In the case of oxides other than those listed (i.e. other than those of the alkaline earths) data by Jentsch[14] show lack of agreement in most instances between β(calc.) and β(obs.). In every such case β(obs.) is less than β(calc.), and this statement holds for any of the above-described methods of calculating the constant A. High values of the valence, according to Eq. (11), lessen the discrepancy but do not eliminate it entirely. It is possible that this lack of agreement may not mean a breaking down of the formula, but that it is due to secondary effects. Thus, as Richardson has pointed out,[15] in the variation of ψ with temperature a term involving the specific heat of electricity was neglected which may not be justifiable when dealing with most poor conductors. Furthermore in the case of emission from a poor conductor like an oxide, the latter may find difficulty in replenishing rapidly enough the supply of free electrons near its surface. Thus the condition for current saturation might be hard to fulfill in the case of emission from high resistance material. Difficulty in attaining saturation is common experience in such investigations.

The fact that the discordant oxides have melting points close to the temperature range of measurement, while the alkaline earth oxides for which agreement is excellent have melting points about 1000°C higher than the region of observation, suggests the possibility that the process of vaporization may affect the magnitude of thermionic currents. Since the rate of vaporization follows a law similar to that of thermionic emission, no change in form of the thermionic equation might be expected if this be true, but its constants would be altered. If vaporization should facilitate electron emission this would tend to explain the results observed.

The state of poor conductors like the oxides, particularly as concerns both conductivity and thermionic investigations, is often apparently an extremely sensitive function of temperature, applied potential, and other physical conditions, and it is frequently difficult to distinguish between reversible and irreversible changes which occur. Under these circum-

[12] Waterman, loc. cit.[5] p. 262, Eq. (7)
[13] Horton, Phil. Mag. **11**, 505 (1906)
[14] Jentsch, Ann. der Phys. **27**, 129 (1908)
[15] Richardson, Electron Theory of Matter, p. 452-454

stances it becomes questionable in many cases to expect that any simple theoretical expression may be verified until the nature of such peculiarities is more fully understood.

Calculation of Free Electron Concentrations Within Conductors

To summarize the above, in Eqs. (4), (5), and (10) we have different expressions for the concentration of free electrons in equilibrium with a hot body. Eq. (4) yields a value of A consistent with experiment provided n_B, the concentration of internal free electrons, is of the order of magnitude of the atomic concentration within the substance. Such a magnitude is difficult to reconcile with other phenomena, notably the contribution of the electrons to the specific heat. According to Eq. (5) the concentration of internal electrons does not enter into the problem, the work function only being characteristic of the hot body. For the reasons above stated this appears merely to be an approximation to the truth. It has been shown that Eq. (9) is fully as satisfactory as Eq. (3) in accounting for the thermionic current. Assuming the validity of the hypothesis underlying Eqs. (8) and (9), it is therefore possible to calculate the concentration of free electrons within a conductor from Eq. (7). These free electron concentrations are given in Table II, which includes the variation of the concentration with temperature for Na. The reason that the concentration in metals increases with decreasing temperature lies in the fact that conductivity considerations in the case of metals require that ψ_0 be slightly greater than ϕ_0, the energy of dissociation $(\phi_0-\psi_0)$ of a bound into a free electron being negative.[5] For poor conductors similar considerations require that $\phi_0-\psi_0$ be positive and large. The valence ν in the atomic reaction liberating electrons is taken to be unity for all substances listed.

Table II

Conductor	N (per cc)	T (°K)	b (°K)	n (per cc)
Na	2.6×10^{22}	273	−26*	$1.0\times(10)^{17}$
		173		$1.5\times(10)^{17}$
		100		$2.6\times(10)^{17}$
		50		$5.6\times(10)^{17}$
		20		$2.4\times(10)^{18}$
		10		$1.5\times(10)^{19}$
Cu	8.5	273	—†	$1.7\times(10)^{17}$
Pt	6.6	273	—†	$1.5\times(10)^{17}$
Ag	5.9	273	—†	$1.4\times(10)^{17}$
Pb	3.3	273	—†	$1.1\times(10)^{17}$
Bi	2.9	273	—†	$1.0\times(10)^{17}$
MoS_2	1.8	273	+3000*	$1.3\times(10)^{13}$
		773	"	$7.4\times(10)^{14}$
CaO	3.3	1273	+21,000*	$2.3\times(10)^{9}$

*from conductivity data; † assumed negligible.

The free electron concentration thus appears lower than most estimates given hitherto. Such an order of magnitude however seems demanded by the study of the specific heats of good conductors. It will be observed that the mean concentration in metals at ordinary temperatures, namely about 10^{17} per cc, gives an entirely negligible electronic contribution to the specific heat of metals, and in the case of Na, which is typical, the specific heat contribution by free electrons remains negligible down to very low temperatures, being 0.0024 at 50°K out of a total of 3.50 cal. per gm atom, the atomic heat of Na at this temperature. Even at 10°K this contribution would only be 0.028. This result appears to be of especial interest and significance, in that these concentrations, which are computed directly from the theoretical expression, remove what has always been a serious obstacle to the progress of a free electron theory.

Concentrations of this order of magnitude necessitate large mean free paths according to the usual simple classical conductivity theory. In the case of Na the electronic mean free path, as calculated from its specific resistance by this method, would be 0.014 cm at 0°C. On the other hand it becomes increasingly probable (cf. results of Ramsauer[16]) that such a mean free path is possible. Bridgman[17] has already concluded that the mean free path in metals is at least 100 or 1000 times as long as the mean distance between atoms. In fact a long mean free path, coupled with the fact pointed out by Eldridge[18] that an electric current is carried by the free electrons whose thermal motions have a component perpendicular to the current flow, may have a share in explaining the abnormally high resistance of thin films.

In conclusion, the hypothesis that in conductors a free electron concentration exists which is determined by chemical equilibrium considerations, yields a value of this concentration as a function of temperature which is competent to explain the observed temperature variation both of electrical conductivity and of thermionic saturation current densities. Furthermore the computed numerical values of these free electron concentrations are not in conflict with a most important requirement of any free electron theory, namely that the equi-partition value for the specific heat of the free electrons be negligible.

SLOANE PHYSICS LABORATORY,
YALE UNIVERSITY,
April 15, 1924

[16] Ramsauer, Ann. der Phys. **66**, 546 (1922)
[17] Bridgman, Nat. Acad. Sci. Proc. **7**, 299 (1921)
[18] Eldridge, Phys. Rev. **21**, 2, 131 (1923)

CHANGE OF CONDUCTANCE OF SELENIUM DUE TO ELECTRONIC BOMBARDMENT

By Ralph de Laer Kronig

Abstract

Effect of electronic bombardment on the conductance of selenium.—A low-resistance selenium cell about 1 inch square made by pressing Se between twenty double turns of fine Pt wire, was mounted in a vacuum tube with a grid and an oxide-coated, equipotential, electrically heated cathode and subjected to bombardment by electron currents up to 400μ-amp. at potentials up to 95 volts. The increase of conductance ΔC with current I and with voltage V, is found to agree well with the theoretical formula $\Delta C/C=\sqrt{1+kIV}-1$, derived on the basis of the *free electron theory of conduction*, assuming that the impinging electrons increase the number of free electrons throughout the whole volume of the Se by ionizations proportional both to the number of electrons and to the energy of each, i.e. to IV. The constant k is the ratio of the ions created per watt to the number of ions produced spontaneously per sec. and comes out 98.5.

Theory

According to the electron theory of metallic conduction the current is carried by the free electrons in the conductor, and the conductance is directly proportional to the number n of free electrons per unit volume. After the conductor has been left sufficiently long under unchanged external conditions, a state of equilibrium must exist between the un-ionized atoms, the ionized atoms, and the free electrons; i.e. the number of free electrons produced per unit volume every second must be equal to the number of free electrons recombining per second per unit volume.

In the case of selenium only a very small fraction of the atoms ever becomes ionized, and the number of un-ionized atoms per unit volume can therefore be regarded as constant. Let n be the number of ionized atoms per unit volume, and therefore the number of free electrons per unit volume. The chance of a recombination is proportional to n^2, and therefore the number of recombinations per unit volume per second may be put equal to a_1n^2.

In the case here dealt with, of selenium subjected to bombardment by electrons, in addition to the ions produced by thermal agitation and by the constant illumination from the filament, whose number per unit volume per second we shall denote by a_2, we may suppose ions are pro-

duced by the impinging electrons uniformly throughout a surface layer of depth d. Let the conductance of this surface layer be changed by the bombardment from C_0 to C_1. It is reasonable to assume that the number of ions produced is proportional to the number of impinging electrons, i.e. to the current I falling on the cell, in future simply called the plate current, and also proportional to the energy of each electron, i.e. to the difference of potential V through which it has fallen and which we shall speak of as the plate potential, therefore proportional to the product IV. For equilibrium in the surface layer we must have, therefore,

$$a_1 n^2 = a_2 + a_3 IV$$

or

$$n = \sqrt{(a_2 + a_3 IV)/a_1} .$$

Therefore the conductance of this layer is

$$C_1 = \sqrt{b_1 + b_2 IV}$$

where b_1 and b_2 are constants. When no electrons are falling on the cell other conditions remaining unaltered, the conductance of this layer is

$$C_0 = \sqrt{b_1} ,$$

so that

$$C_1 = \sqrt{C_0^2 + b_2 IV} . \tag{1}$$

For the ratio of the change in conductance ΔC to the original conductance of the whole cell C, we have

$$\frac{\Delta C}{C} = \frac{C_1 - C_0}{C} = \sqrt{\frac{C_0^2}{C^2} + k_2 IV} - \frac{C_0}{C} = \sqrt{k_1^2 + k_2 IV} - k_1 \tag{2}$$

where $k_1 = C_0/C$ and k_2 is a constant independent of I and V.

If the product IV is sufficiently small, we may write,

$$\Delta C/C = \tfrac{1}{2}(k_2/k_1) IV . \tag{3}$$

Apparatus

The selenium cell used had a dark resistance of about 3500 ohms. It was made by winding 20 turns of each of two platinum wires, about .01 cm in diameter, side by side on a slab of insulating material about 2.5 cm square, and pressing the selenium between the wires on one side. A shellac coating on the other side prevented the wires from touching. The source of bombarding electrons consisted of a heating strip of platinum, .025 mm thick, 3 cm long, and .1 cm wide and an outer platinum strip, .05 mm thick, insulated from the inner one by a sheet of mica. The outer

strip was coated with a half and half mixture of barium oxide and strontium oxide and connected to one terminal of the inner one, thus forming an equipotential emitting surface when a heating current flowed through the inner strip. Cell *C* and filament *F* together with a grid *G* were mounted on the lead-in wires passing through the ground glass stopper of the glass tube (Fig. 1). The grid was 3.5 cm on edge, having approximately 2 wires per cm. The tube was kept evacuated by means of a rotating mercury pump and a mercury diffusion pump in series, a mechanical pump serving as fore-pump. The mercury vapor was frozen out in a trap surrounded by carbon dioxide snow or a mixture of carbon dioxide snow and acetone, the last traces being removed by a piece of gold foil inserted in the exit of the vacuum tube. This was found to be very important, not only because the vapor would give rise to ionization at high plate

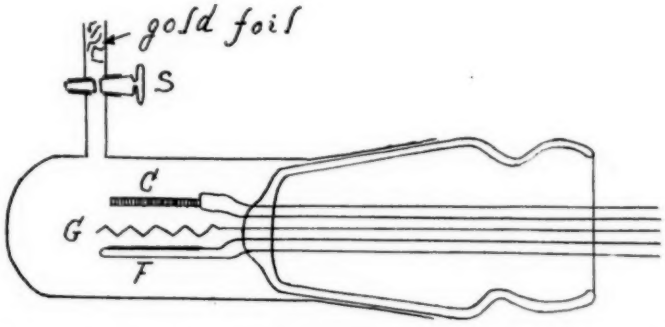

Fig. 1. Diagram of vacuum tube containing selenium cell *C*, grid *G* and Wehnelt cathode *F*.

potentials, but also because it affects the surface of the selenium, greatly increasing its conductivity. A stop cock *S* was therefore inserted in the tube leading from the vacuum chamber and was opened only after the cooling mixture had had sufficient time to remove the mercury vapor. The pressure in the tube was read on a MacLeod gage; it ranged around 10^{-5} mm. The outside light was excluded by means of a black cloth so that the cell was exposed only to the constant illumination from the filament. The voltage between cell and filament could be varied from a potentiometer connected to a storage battery. An adjustable voltage on the grid regulated the plate current without altering the filament current.

The cell was inserted in one arm of a Wheatstone bridge in parallel with a known resistance *r*, whose midpoint was connected through the plate voltage supply to the equipotential filament, a galvanometer or microammeter *A* serving to read the plate current (Fig. 2.) and a volt-

meter V indicating the plate potential. The resistance r was necessary in order that none of the plate current might flow through the bridge galvanometer. Furthermore, in order that the electrons striking the cell should all have the same velocity, say within .1 volt, the drop across the cell due to the bridge current was kept as small as possible, never exceeding .1 volt. It was on this account that a low-resistance cell was used, because with so small a voltage no precision could otherwise have been obtained in the resistance measurements.

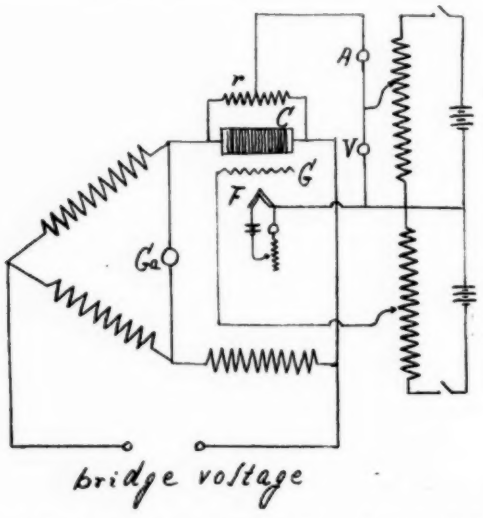

Fig. 2. Electrical circuits used in measuring the resistance of the selenium cell, and the plate current and the plate and grid voltages.

Results

The resistance of the combination C and r in parallel was measured for different constant plate currents, varying the plate voltage up to 100 volts, and from this resistance X' the resistance X of the cell was computed, and then the ratio $\Delta C/C$ plotted as a function of the plate voltage V. Fig. 3 shows three characteristic curves for constant plate currents of 50, 150, and 400 micro-amperes respectively. Similarly Fig. 4 shows three characteristic curves obtained by keeping the plate voltage V constant and finding the ratio $\Delta C/C$ as a function of the plate current I, the curves corresponding to values of V of 20, 50, and 80 volts respectively. From the curves it is evident that for sufficiently small values of I and V the curves are practically straight lines and can therefore be represented by a formula of the type of Eq. (3),

$$\Delta C/C = \tfrac{1}{2} kIV$$

For larger values of plate currents and plate voltages the values found for $\Delta C/C$ are in good agreement with the formula

$$\Delta C/C = \sqrt{1+kIV} - 1 \qquad (4)$$

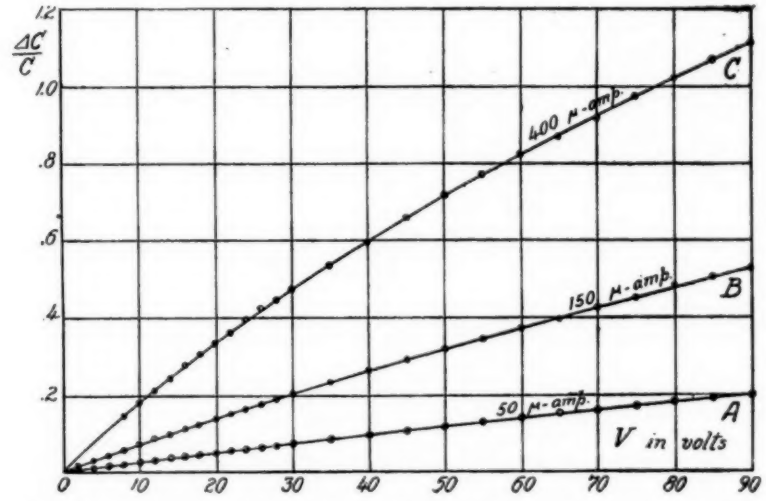

Fig. 3. Ratio $\Delta C/C$ as function of the plate voltage for constant plate currents of 50(A), 150(B) and 400(C)μ-amp.

which is identical with Eq. (2), if $k_1=1$. For if the values of $\Delta C/C$, I, and V are used to compute k, it comes out constant over the range of observations, (see Table I). We therefore find that k_1 of the theoretical

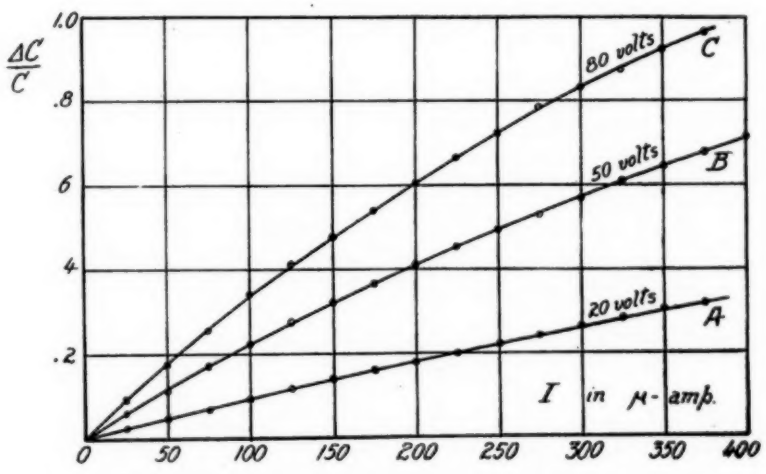

Fig. 4. Ratio $\Delta C/C$ as function of the plate current for constant plate voltages of 20(A), 50(B), 80(C) volts.

derivation is equal to unity. Therefore $C_0 = C$, or the original conductance of the affected layer is equal to the original conductance of the whole cell. In other words ionizations due to the impinging electrons take place throughout the entire material at approximately the same rate and do not confine themselves to a surface layer. This is in excellent agreement with observations by Brown,[1] which indicate that the change in conductance due to illumination of selenium spreads from the illuminated regions of the material to the unilluminated portions, sometimes as much as 10 mm. The explanation for this behavior, i.e. the mechanism for the necessary transfer of energy will probably have to be sought in the production of secondary radiation.

TABLE I

V (volts)	I (micro-amperes)	k
20	100	98.7
20	200	98.0
20	300	97.5
50	100	99.3
50	200	99.0
50	300	97.7
80	100	99.3
80	200	98.3
80	300	98.5

The writer wishes to express his gratitude to the department of physics for many helpful suggestions and especially to Professor A. P. Wills for the constant interest taken by him in the progress of this research.

PHOENIX PHYSICAL LABORATORIES,
COLUMBIA UNVERSITY,
May 26, 1924.

[1] F. C Brown, Phys. Rev. **5**, 404 (1914)

COEFFICIENTS OF DIFFUSION OF CERTAIN ALKALI SALT VAPORS IN THE BUNSEN FLAME

By George E. Davis

Abstract

The method used is a modification of one suggested by H. A. Wilson. The ellipsoidal streak of luminous vapor from a salt bead held in the flame, was photographed; then the coefficient of diffusion K was obtained from the equation $K=\frac{1}{2}rv(dx/dr-1)$, where r is the radial distance from the bead to a point on the boundary of the streak, x is the vertical distance of the point above the bead, and v is the velocity of the flame gases. A Meeker gas burner of special design was used, supplied with air under pressure and with intermittent puffs of salt spray to enable v to be determined. The flame temperature, determined by Féry's method, was 1436°C. The means of about ten determinations of K for each salt are: *lithium chloride*, 12.4; *lithium sulphate*, 9.7; *sodium chloride*, 19.0; *sodium sulphate*, 16.4; *potassium sulphate*, 11.1; *rubidium sulphate*, 11.7; *caesium sulphate*, 8.4. These are probably accurate to about 10 per cent. The chlorides of both Li and Na diffuse faster than the sulphates. Except in the case of the lithium salts, for which the observed values are relatively too low, the diffusion coefficients increase roughly as the reciprocals of the square roots of the weights of the metal atoms, also roughly as the reciprocals of the square roots of the molecular weights. If the free metal atom of the salt is the source of luminosity, as some other experimenters have concluded, it is probable that it is free only a fraction of the time, the fraction varying with the salt and the alkali.

IN 1912, H. A. Wilson published the results of some measurements of the coefficients of diffusion of alkali salt vapors in the Bunsen flame.[1] Some time later, while working under the direction of Prof. Wilson, the author made some new measurements of the coefficients of diffusion of a number of salts, including certain salts of the alkali metals. In these determinations, the data obtained for any one salt were not enough to make the result very reliable. More recently a new series of measurements has been carried out. A photographic method suggested by Prof. Wilson is used, with certain modifications which simplify the calculations and increase the accuracy of the results.

Theory

The streak of luminous vapor extending upward from a bead of salt held in the non-luminous Bunsen flame appears to the eye to have a rather well defined boundary. This apparent boundary is the locus of

[1] H. A. Wilson, Phil. Mag., July 1912

points at which the concentration of salt vapor is equal to the minimum value for visible luminosity. Therefore the concentration is the same at all points on this boundary. In the paper published by Wilson in 1912,[1] it is shown that the concentration C at any given point at a distance r from the salt bead is given by

$$C=\frac{q}{4\pi Kr}e^{(v/2K)(x-r)}, \tag{1}$$

where q is the amount of salt evaporating in unit time, K is the coefficient of diffusion, v is the velocity of the flame gases, and x is the distance to the given point measured along the axis of the flame. From the graph of this equation we find that any surface of constant concentration resembles an ellipsoid of revolution with its major axis lying along the axis of the flame. Hence the apparent boundary of the luminous streak of salt vapor should be such a surface. Expressing r in terms of x and y, in Eq. (1), differentiating with respect to x, and setting dy/dx equal to zero, gives

$$K=\frac{vr_m}{2x_m}(r_m-x_m), \tag{2}$$

where r_m and x_m are the values of r and x where the horizontal diameter of the streak is maximum. Wilson observed x_m and the maximum diameter of the streak directly and computed K by Eq. (2). The principal objection to this method is that the maximum diameter of the luminous streak is very difficult to observe and locate definitely by direct observation.

In the experiments performed under the direction of Prof. Wilson, referred to above, a different equation was used. From Eq. (1) we may obtain

$$K=\frac{v(x_2-x_1-r_2+r_1)}{2\log(r_2/r_1)}, \tag{3}$$

where the subscripts refer to any two given points on the apparent boundary of the luminous streak. The values of x and r were obtained by photographing the luminous streak and taking measurements on the photographic plate. An average value of K was found from the results for several pairs of points.

In the present measurements a still different equation was used. Differentiating Eq. (1) with respect to r, we get

$$K=\frac{rv}{2}\left(\frac{dx}{dr}-1\right), \tag{4}$$

which holds for points on the apparent boundary of the luminous streak. Values of x and r were obtained from the photographic plate, as before. These values were plotted and the slope dx/dr obtained from the resulting curve for a number of points corresponding to a chosen series of values of r. For such a series of points, we obtain from Eq. (4)

$$K = \frac{v}{2\,(1/r)_{av}} \left[\left(\frac{dx}{dr}\right)_{av} - 1\right].$$

Comparing this with Eq. (4), it is seen that for the average slope there is a corresponding value of r, say $\bar{r}$, equal to the reciprocal of the average value of $1/r$ and independent of K and v. $\bar{r}$ and the average value of

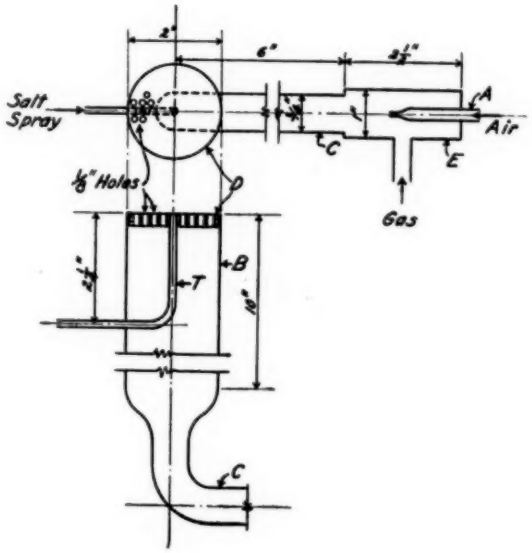

FIG. 1. Gas burner

dx/dr were substituted in Eq. (4) for r and dx/dr, respectively. This method of averaging values makes only one substitution in the equation necessary. The same value of $\bar{r}$ can be used for a number of different curves.

EXPERIMENTAL METHOD

In order to obtain a large gas flame, a Meeker type of burner of special construction was used (Fig. 1). A large number of holes were drilled close together in an iron disk D. This formed the burner top. B is a vertical brass tube, joined to a smaller horizontal tube C. The chamber E consists of an iron tee connected with the air intake tube A by means of

reducing fittings. Compressed air was used. Tube *A* was arranged as shown in order that the air pressure might not force the gas back. The long length of tubes *B* and *C* allowed the gas and air to mix thoroughly before being burned. Shorter lengths, with baffle plates as mixers, were tried, but were found impractical because of the tendency of the flame to "strike back." A tube (*R*, Fig. 3, not shown in Fig. 1) 5 in. long, diameter 2.5 in., was supported so as to be coaxial with tube *B* and extend about .75 in. above the burner top. This tube helped to steady the flame. The flame appeared to be perfectly uniform and did not taper much for a distance of several centimeters above the burner. The small tube *T* was used to introduce puffs of salt spray into the flame in order to measure the velocity of the flame gases.

To obtain these puffs of salt spray, a special type of valve (Fig. 2) was designed. *C* is a cylindrical chamber made by drilling out a solid brass cylinder. *D* is an iron disk screwed onto the end of a shaft *S* fitted with sleeve and set screws (not shown) for connecting with the shaft of an

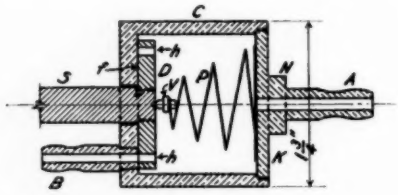

FIG. 2. Valve for giving puffs of salt spray

electric motor. Face *f* of disk *D* is machined smooth and made accurately perpendicular to the axis of *S*. The inner face *f* of the cylinder end also is turned smooth and is held against *D* by means of spring *P* which carries a small pivot *V*. This arrangement allows *S* and *D* to turn freely without imparting much torque to *C*. Near the circumference of *D* are drilled 8 small, evenly spaced holes *h*, so placed as to come in turn into momentary alignment with the hole through nipple *B*, as *D* rotates. The opposite end of chamber *C* is closed by a brass disk *K* which carries a nut *N* and a nipple *A*, all machined from the same piece. When in use, the barrel *C* is supported by a rubber band placed around it, which keeps *C* from turning yet takes up slight movements due to lack of perfect alignment of shaft *S* with the motor shaft. Salt spray entering through *A* passes out at *B* in a series of puffs as disk *D* rotates. These puffs are led into the flame through tube *T* of Fig. 1.

The general arrangement of apparatus is shown in Fig. 3. *W* is a platinum wire with a very small loop at the end, holding a salt bead *B* in the

center of flame F as seen from above. The section of W heated by the flame was adjusted so that only the salt bead showed on the photographic plate when an exposure was taken with camera C. S is a vertical steel millimeter scale with its edge nearest B located in the plane which passes through B perpendicular to the axis of the camera lens. A Bausch and Lomb achromatic lens was used, diameter of full aperture $1\frac{5}{8}$ in. It gave an image free from distortion, using full aperture. Panchromatic plates were used. Adjustment was made for nearly equal size of object and image.

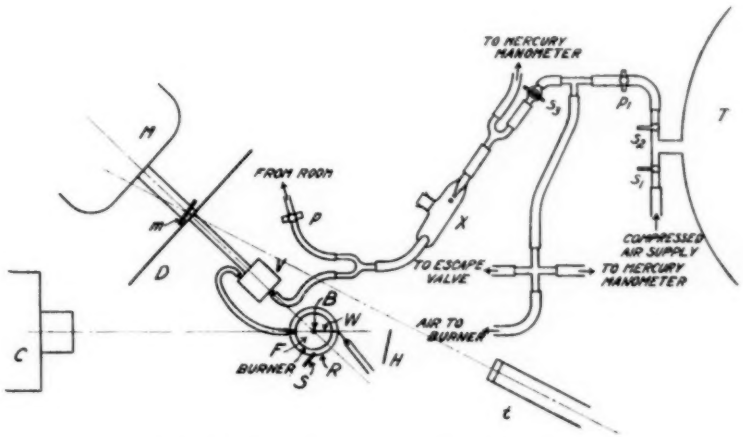

FIG. 3. General arrangement of apparatus

Connections were made with the compressed air line as shown. A tank T of about 12 cu. ft capacity helped to steady the flow. The pressure of the air going to the burner was regulated by a special type of escape valve, the essential parts of which are shown in Fig. 4. This consisted of a brass tube A working easily but snugly inside a second tube B. Melted lead poured into A closed the upper end and furnished part of the necessary weight. Adjustment of weight was effected by adding lead shot. Vaseline was used to lubricate the movable tube and prevent the escape of air between the tubes. Tube A had a number of very small holes drilled through it in two vertical rows located opposite each other. The action of the valve as a pressure regulator is obvious. It kept the pressure constant within a range of 1 or 2 cm of mercury. It was supplemented by hand regulation when necessary, using stopcock S_1 and pinchcock p_1 (Fig. 3). As a rule, the air pressure was adjusted to about 30 cm of mercury. Then the gas pressure was adjusted until the inner cones of the flame were just below the height at which they flared up. These condi-

tions were kept as nearly constant as possible in all the determinations. Mercury U-tube manometers were connected as shown. A tilted U-tube with water in it indicated the pressure in the gas line (not shown).

The velocity of the flame gases was determined as follows. Compressed air was let into atomizer X by opening stopcock S_3. Salt spray, from a solution of sodium chloride in X, passed partly to valve V (previously described) and partly through a tube leading out of the room. The shaft of V was connected directly to the shaft of an electric motor M. The motor shaft also carried a 10 in. metal disk D with 8 evenly spaced radial slits cut in it. When D and the disk of valve V were rotated, 8 puffs of salt spray were introduced into flame F per revolution. Viewed through the 8 slits in D, conveniently by means of mirror m, these puffs appeared to be stationary in the flame. The distance between puffs was measured by use of a cathetometer with telescope t. H is a screen for t. Settings on

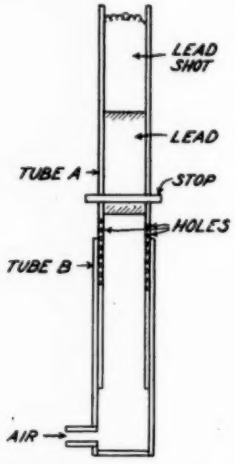

FIG. 4. Constant pressure air valve

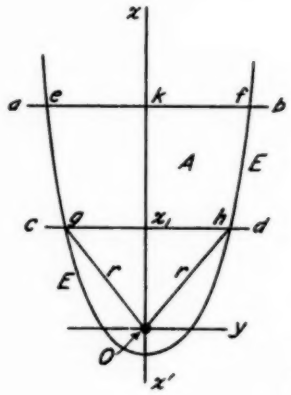

FIG. 5

the puffs were made where their diameters were largest. Three puffs were visible at one time. Multiplying the speed of the motor in r.p.s., determined with a revolution counter, by 8 times the distance between puffs, gave the velocity of the flame gases.

The temperature of the flame was found by Féry's method,[2] using the carbon filament of an incandescent lamp as the body whose temperature is adjusted to that of the flame. Adjustment was made for reversal of the sodium D-line. The filament temperature was measured by means of a Leeds and Northrup optical pyrometer. As a check upon results, readings also were taken on the flame from an ordinary Bunsen burner.

[2] Féry, Comptes Rendus, **137**, 909-912 (Nov. 30, 1903)

These latter readings fell above the scale limit of the pyrometer but could be estimated without much error.

All the experiments were performed in a darkroom. The waste flame gases and salt vapor were led out of the room by a flue made of 6 in. stove pipe, supported with the lower end directly over the flame.

Several photographic plates (on the average eight) were exposed during each run, without allowing the flame adjustment to change. Usually two or more different salts were represented in each such group of exposures. This made possible a more accurate comparison of the coefficients of diffusion of the various salt vapors.

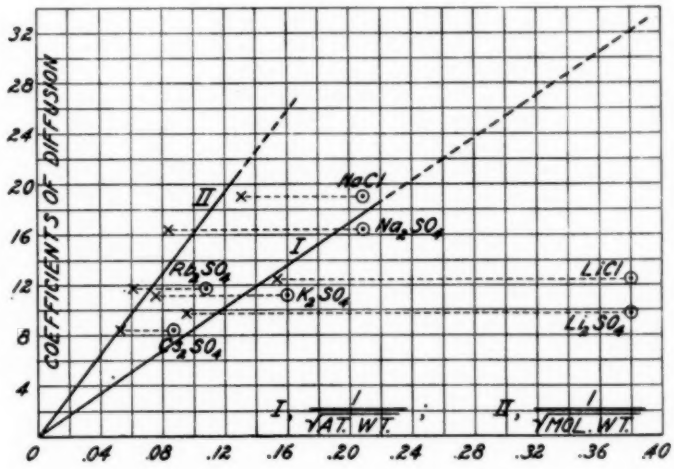

Fig. 6. Coefficient of diffusion of salt vapor as function of reciprocal of square root of weight of metal atom (I), and of molecular weight (II).

The method of marking the plates was as follows. Referring to Fig. 5, the center of the image *O* of the salt bead was first marked with a needle prick. Then a line *ab* was cut across the image *A* of the luminous streak as far from *O* as the distinctness of the apparent boundary *EE* of the streak would allow. This line was cut as nearly perpendicular to the axis *xx'* as could be judged before the axis was located. The points *e* and *f* where the line met the boundary *EE* were then marked, *ef* carefully bisected, and a line *xx'* cut through the middle point *k* and the center of the bead. Then the plate was put on the table of a dividing engine and from 7 to 15 evenly spaced lines like *cd* cut 2 or 3 mm apart, perpendicular to *xx'*, above *O*. The points where these lines met the boundary *EE* were then marked. In many cases the image *A* faded out quite gradually at the edge, so that it was somewhat difficult to locate these points. In

locating them, some consideration was given to the fact that the theory of diffusion indicates that the points should fall on a smooth curve on either side.

A different method of locating points of equal intensity of blackening was tried, using a microphotometer of the form described by Meggers and Foote,[3] and considering points near the apparent boundary *EE*. This method seemed to promise more accurate results than could be obtained by the method of marking just described. However, the results actually obtained were often inconsistent and unsatisfactory, so the method was not used.

To measure the plates, they were placed on a small table carried by the movable table of a dividing engine and viewed by means of a fixed microscope mounted above. Distances could be read accurately to .001 cm. The distance (as *gh*, Fig. 5) from edge to edge of the image of the luminous streak was measured along each of the cross-lines and the corresponding radial distance r calculated. The distance representing unit length on the plate was obtained by taking a series of readings on the image of a scale appearing on the plate along the image of the luminous streak. Incidentally it was found that interposing a sheet of reddish-brown glass between illuminator and plate made the images on the plate stand out much more clearly than when illuminated by white light.

Next, x was plotted against r on accurately ruled paper of large size. The curves were from 45 to 80 cm in length. They were very carefully drawn and the plotted points usually fell on, or very close to, a smooth curve. The coefficient of diffusion of the salt vapor was obtained from the curve by the method described under "Theory." The values of the slope dx/dr were found by laying a straight-edge tangent to the curve at the chosen points. This could be done with a fair degree of accuracy, since the curves were very nearly straight in every case.

It is assumed in the mathematical theory that all the stream lines of the moving medium are parallel. Therefore it is important to determine whether the tapering of the flame and the distortion of the stream lines near the salt bead have an appreciable effect upon the results. As a test, a group of 25 plates giving long x, r curves were chosen at random. Two determinations of the diffusion coefficient K were made from each curve, one from the upper 5 values of r in the series already chosen, and one from the lower 5. Then the average K obtained from the upper points was compared with that from the lower. As a further test, two groups of 4 plates each were chosen, in each group the kind of salt and the flame

[3] Meggers and Foote, Scientific Papers Bur. Standards, No. 385 (June 30, 1920)

velocity being the same for all the plates. From the average dimensions in each group, a curve was plotted to represent the apparent boundary of the image of the streak of vapor. This curve was compared with a theoretical curve which had arbitrarily been made to coincide with it at two points. The results are given below.

Experimental Results

In Table I is shown a typical set of values of dx/dr obtained from one of the x, r curves. In Table II are given the results obtained from the individual photographic plates, with their averages. The results obtained by H. A. Wilson are shown for comparison. The particular salts used in his determinations are not mentioned in his article, so are not specified here.

In Table III are shown the temperature readings at three different points in the flame, under different conditions as to pressure of air supply to the burner. Since the air pressure determines the velocity of the flame gases, the measurements show that the temperature was not appreciably affected by velocity variations much greater than those occuring during the determinations of the diffusion coefficients. The temperature is the same at all three points in the flame, within the limits of accuracy of the measurements. These points are so located as to give a fair indication of the temperature throughout the region in which the coefficients of diffusion were determined. Therefore we conclude that the results were not appreciably affected by variations in temperature from point to point in the flame. The average value, 1436°, may be a little lower than the true temperature, since the carbon filament used in the measurements was not a perfect black-body radiator. The true temperature may be from 45° to 90° higher.[4] However, the value 1850° for a Bunsen flame of the ordinary type agrees well with the most reliable values found by other observers.

The gas burned was manufactured from bituminous coal and may be described as a mixture of coal gas and "water gas." Its composition varied somewhat from day to day, as shown by analyses made by the department of chemistry. These analyses showed the following content, in volume per cent: CO, 11 to 16; CO_2, 4 to 7; O_2, 1.5 to 6; unsaturated hydrocarbons, 6 to 17; CH_4, 11 to 16.5; H_2, 18 to 21; N_2, 27 to 32.

In the first test made to determine whether the values found for the coefficients of diffusion K were affected appreciably by lack of parallelism of the stream lines of the flame gases, the average value of K for the

[4] Burgess and LeChatelier, "Measurement of High Temperatures"; observations by Burgess, pp. 340, 341.

upper 5 values of r came out 4.8 per cent higher than the average for the lower 5. In the second test, comparing theoretical shapes of streaks of salt vapor with shapes experimentally determined, a small departure of the latter from the former was found in each case. These tests seemed to indicate an appreciable but rather small effect on the values found for K, due to lack of parallelism of the stream lines.

TABLE I

Typical set of results

r	x	dx/dr*
.800	.614	1.068
1.000	.825	1.053
1.200	1.035	1.040
1.400	1.243	1.035
1.600	1.450	1.033
1.800	1.656	1.030
2.000	1.862	1.028
2.200	2.068	1.023
2.400	2.273	1.020
2.600	2.477	1.018
2.800	2.680	1.015
3.000	2.884	1.013
		Average $dx/dr = 1.031$

Reciprocal of average value of $1/r = 1.616$. Velocity of flame gases = 328 cm/sec. Coefficient of diffusion computed by Eq. (4) = 8.2.

* These values were obtained from a plotted curve, not computed from the values of r and x.

TABLE II

Coefficients of diffusion

	LiCl	Li_2SO_4	NaCl	Na_2SO_4	K_2SO_4	Rb_2SO_4	Cs_2SO_4
	12.7	12.0	18.8	16.3	12.9(3)	11.7	8.2
	14.5	12.9	22.6	18.7	17.1	11.1	8.5
	15.0(3)	10.8	15.8(4)	14.8	14.2	13.7	8.7
	16.6	10.2	20.8	20.9	12.5	11.4	9.0
	11.9	11.1	22.9(4)	21.8	8.6	11.3	7.8
	10.9(4)	6.7	19.5	15.7	9.5	13.1(3)	9.6
	9.3	8.4	14.0	14.8	8.7(4)	10.6(4)	7.5
	15.3	9.6	19.2	13.3	9.1(4)	12.0	8.0
	10.8	8.5	18.6	15.4	6.3(4)	10.3(4)	8.0
	11.9	6.5(4)	15.8(4)	13.7			
	7.9	8.9(3)	20.4	14.6			
Weighted average*	12.4	9.7	19.0	16.4	11.1	11.7	8.4
Average deviation	±2.2	±1.6	±2.1	±2.3	±2.8	± .8	± .5
Wilson's results†	(Li)14.5		(Na)11.5		(K)4.7	(Rb)4.7	(Cs)4.4

* Scale of weights, 3 to 5; the weight is 5 unless otherwise indicated by figures in brackets; weighted on basis of excellence of photographic plate and apparent reliability of data.

† The particular kind of alkali salt examined in each case is not stated by Wilson.

TABLE III

Flame temperatures, by Féry's method

Position in flame	Pressure of air supply	Temperature
Lower center of flame, just above salt bead	18.3 cm Hg	1435°C
	30.8	1440
	31.1	1433
.8 cm from boundary of flame, just above level of salt bead	30.9	1426
Center of flame, about 2.5 cm above salt bead	30.9	1444
	49.0	1440
		Av. temp. = 1436°

Reading on Bunsen flame of ordinary type, with full air supply = 1850°.

DISCUSSION

The coefficients of diffusion found in these measurements (Table II) do not agree well with those found by Wilson. It is believed that the results obtained in the present measurements are the more reliable for reasons given under Theory. A difference in composition of the gases burned in the two flames, or a difference in flame temperatures, may partly account for the disagreements. The above results are probably accurate to about 10 per cent.

It has been shown by Smithells, Dawson and H. A. Wilson[5] that the light emitted by a flame containing vapor of an alkali salt is probably due to atoms of the alkali metal. They pointed out that a metal atom might lose one or more electrons under the influence of violent collisions with other atoms in the flame, thus becoming a positive ion; and that this ion, upon recombination with other free electrons in the flame, might be subject to internal vibrations which would give rise to the radiation of light. This theory was strengthened by the results of later experiments by H. A. Wilson.[6] These experiments showed (1) that the luminous vapors of salts in a Bunsen flame are not deflected appreciably by an electric field; (2) that the positive ions present in the luminous vapor can be made to move out of it by an electric field; (3) that the positive ions are not luminous, but can form luminous vapor after recombination. It has also been shown by Bancroft and Weiser[7] that many metallic salts are dissociated in the Bunsen flame, the metal atom being set free.

[5] Smithells, Dawson and H. A. Wilson, Phil. Trans., A **193,** 89 (1899)
[6] H. A. Wilson Phil. Trans., A **216,** 63-90 (1915)
[7] Bancroft and Weiser, Journal of Phys. Chem., 1914.

If, as this evidence indicates, the metal atom of the vapor is the source of light, we might expect the coefficient of diffusion, determined from the shape of the luminous streak of vapor from a salt bead, to be the coefficient of diffusion of the free metal atoms. But it is likely that these atoms, because of their tendency to combine with other ions (radicals) in the flame, are free only part of the time. Therefore the average rate of diffusion for any particular atom probably is less than the rate we should expect if the atom were free during the whole of its journey. It therefore seems reasonable to assume that our observed values for the coefficients of diffusion are lower than the values for free metal atoms.

It is interesting to examine our results in the light of two different assumptions: (1) That the metal atoms were free during the entire time spent in diffusing; (2) that the molecules of salt vapor remained undissociated practically the entire time. In the first case, kinetic theory indicates that the observed diffusion coefficients should vary directly as the reciprocals of the square roots of the weights of the metal atoms. In Fig. 6, curve I, these quantities are plotted against each other (points marked with circles). In the second case, the observed coefficients should vary as the reciprocals of the square roots of the molecular weights of the salts. Plotting these quantities, we obtain curve II (points located by crosses). In either case, the values for the sodium, potassium, rubidium and caesium salts locate, roughly, a straight line passing through the origin, as we should expect. On the other hand, lithium chloride and lithium sulphate do not fall near the line located by the other salts in either case, their observed coefficients of diffusion being only roughly 1/3 and 1/2 the values indicated by curves I and II, respectively. The degree of accuracy with which each curve is located by its set of five points is about the same for the two cases. Again, the positions of the two sodium salts relative to a straight line drawn between them from the origin, depend on whether assumption (1) or assumption (2) is made. The same is true of the two lithium salts. These facts suggest that neither assumption is correct, but that the metal atoms are free part of the time and combined with other particles the rest of the time. As to the two lithium salts, their relatively low rates of diffusion may be due to diffusion of the particles in groups rather than singly.

Partial dependence of the rate of diffusion upon the negative radical of the salt is shown by the observed differences between the chlorides and the sulphates of lithium and sodium. In each case the chloride diffuses more rapidly than the sulphate of the same metal. The differences seem too large to be attributed to experimental errors alone. Referring to

Table II, we see that 4 values obtained for lithium chloride are higher than the highest value for lithium sulphate; while 7 values for sodium sulphate are lower than the lowest value for sodium chloride except one.

The various chemical and physical phenomena which occur in the flame are not yet well enough understood to warrant drawing more definite conclusions.

It is a pleasure to the writer to acknowledge indebtedness to Professor L. B. Spinney for his kindness in securing the necessary equipment for this work; also to Professor G. E. Thompson and especially to Professor J. W. Woodrow, for many helpful suggestions.

LABORATORY OF PHYSICS,
IOWA STATE COLLEGE,
AMES, IOWA.
March 8, 1924

AN ELECTROMETER METHOD FOR MEASURING DIELECTRIC CONSTANTS OF LIQUIDS

By A. P. Carman

Abstract

Differential idiostatic electrometer for the measurement of dielectric constants of liquids.—The suspended system includes two cylindrical quadrants C_1 and C_2 suspended so as partly to overlap two fixed cylindrical plates A and B coaxical with C_1 and C_2 and so placed and connected that the torque due to the condenser AC_1 is opposite to that due to the condenser BC_2. To determine the dielectric constant of a liquid, one of the condensers is immersed in it and the other immersed in a liquid or gas of known constant, and the differences of potentials $V_A - V_1$ and $V_2 - V_B$ for zero deflection are read by the potentiometer method. The dielectric constants are then proportional to $(V_A - V_1)^2/(V_2 - V_B)^2 = R_1^2/R_2^2$. This method gave for the *dielectric constant for distilled water* at 25°C with 60 cycle e.m.f. the value 78.07.

THE instrument used is a form of idiostatic electrometer, or electrostatic voltmeter, part or all of which is immersed in a liquid or liquids. The immersed electrometer for measurement of dielectric constants was first used by Silow[1] in 1875, and similar methods have been used by several investigators[2] since then. Two investigators, Perot[3] and Heerwagen[4] employed differential instruments, each using a double quadrant electrometer with the needles on a common stem, of which the lower set of quadrants and needle can be immersed in the liquid. The theory of these two methods is similar to that presented below, but the construction, adjustments and manipulation of the instrument are very different from those of the instrument described in this paper.

The electrometer used is shown in Fig. 1. The plates C_1 and C_2 of the needle system are cylindrical and are suspended to rotate about the vertical axis. The fixed plates A and B are parts of a coaxial external cylinder, placed as shown in the figure to produce a torque on the needle when differences of electrical potential between the needle and the plates are maintained. The shape of the wire frame carrying the needle plates is shown in the figures and, as indicated, either plate C_1 and its opposite fixed plate A or C_2 and B can be immersed in a given liquid by raising

[1] Silow, Ann. der Phys. **156**, 389 (1875)

[2] Cohn and Arons, Ann. der Phys. **33**, 13 (1888); Tereschin, Ann. der Phys. **36**, 792 (1889); Tomaszewsky, Ann. der Phys. **33**, 33 (1888); Rosa, Phil. Mag. **31**, 188 (1891)

[3] Perot, Jour. de Phys. **10**, 149 (1891)

[4] Heerwagen, Wied. Ann. **48**, 35 (1895)

around the plates a beaker containing the liquid. The suspension is by a long phosphor-bronze ribbon which also serves for charging the needle to the desired potential. The needle is damped by means of an attached vane, which dips into a beaker of oil. The deflections of the needle system are read by the standard method of mirror, lamp and scale.

In the actual instrument which was made and used, the plates A, B, C_1 and C_2, were each 2.5 cm square and the needle plates were 6.5 cm apart. The distance between the needle and fixed plate and also their relative

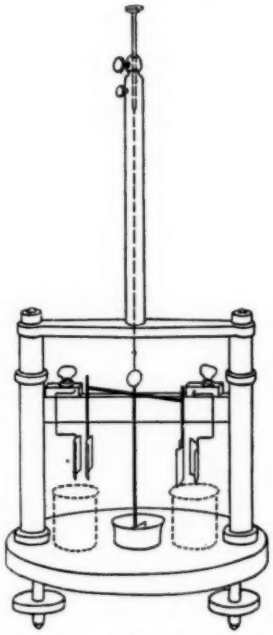

Fig. 1A

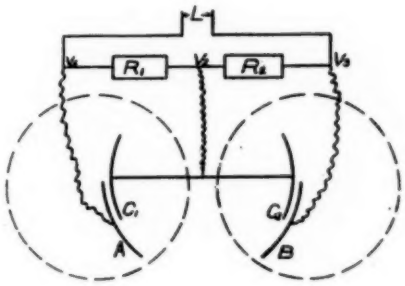

FIG. 1B

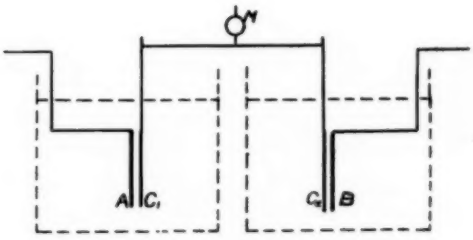

Fig. 1C

Fig. 1A. The electrometer. Fig. 1B. Electrical connections.
Fig. 1C. Section through the plates.

positions can be varied. Other sizes can of course be used and a design is possible in which the pairs of plates are not the same in size on the two sides. Gold plates have been used for liquids which corrode other metals. For some dense liquids, the boyancy is compensated by small lead weights on the horizontal bar of the needle.

The electrical connections are shown in Fig. 1B. The plate A is brought to the potential V_1, the needle plates C_1 and C_2 to V_2, and the plate B to V_3, the potential V_1 being higher than V_2, and V_2 being higher than V_3. This is conveniently done by the usual potentiometer device of tapping off connections from a uniform wire which is stretched zig-zag on a

board of convenient length. In many of our experiments, the potential difference $V_1 - V_3$ was 110 volts, and for conducting liquids this was alternating.

The plates are placed so that the charges on the pairs of plates tend to rotate the needle in opposite directions. Now for equal arm lengths the moments of force are respectively proportional to the squares of the difference of potential $(V_1 - V_2)^2$ and $(V_2 - V_3)^2$, and these in turn are proportional to R_1^2 and R_2^2, where R_1 and R_2 are the resistances of the two parts of the potentiometer wire. But the forces are also proportional to the dielectric constants K_1 and K_2 of the media between the pairs of plates A,C_1 and B,C_2. Hence for the null position of the needle, we have

$$\frac{K_1}{K_2} = C\frac{R_1^2}{R_2^2}.$$

To determine the calibration constant C, the resistances R_{10} and R_{20} for the zero balance when air is the medium between the plates on both sides are found; then if the resistances are R'_1 and R'_2 when the first pair of plates has the medium air and the second pair of plates has the given liquid X as the medium, assuming the dielectric constant of air to be unity, we get

$$C = (R_{20}/R_{10})^2\,; \qquad K_X = (R_2'/R_1')^2 \times (R_{10}/R_{20})^2\,.$$

By immersing the first pair of plates in a medium Y, and the second pair of plates in a medium X, we can calculate the ratio K_X/K_Y of the dielectric constants of the two media from the ratios of the resistances in the calibrating and the comparison balances. This use of a chosen liquid instead of air for reference, has at times decided advantages. Thus in measurements with a substance such as water having a high dielectric constant, we can avoid a large and inconvenient ratio of R_2 and R_1 by using a reference substance with a fairly high dielectric constant, such as acetone or alcohol. The use of a liquid of high dielectric strength for reference instead of air also allows the convenient use of higher potential differences when the consequent increase of sensitivity is desired.

The above instrument is simple in construction, with few parts; its manipulation and adjustments are easily made, and the plates are readily accessible for cleaning, a very important feature in working with most liquids. As the plates on either side can be varied in position independently, the sensitivity and range are easily controlled. As the needle is brought back to zero each time, the actual readings used are readings of resistances which can be made with great accuracy. The immersion method has the advantage of allowing the use of alternating electromotive

forces of a wide range of frequencies. For alternating currents the resistance readings must of course be corrected for reactance. The dielectric constants of semi-conducting liquids can also be satisfactorily measured by this method, particularly if leakage currents are measured and corrections are made for the potential differences caused by leakage.

The above method has been used for determining the dielectric constant of water. The water was doubly distilled in pyrex glass vessels, and the first third and the last third of the distillate were discarded, so that the water used was probably fairly pure. The value of the dielectric constant obtained for this water was 78.07 at 25°C for an alternating e.m.f. of 60 cycles, a value in satisfactory agreement with accepted values. Most of the work which we have been doing with this method has been on the dielectric constants of mixtures of liquids and of emulsions and is not yet ready for publication.

In developing the above method and instrument, the author acknowledges the great help received from Mr. F. E. Null, graduate scholar, who has carried out the experimental work.

LABORATORY OF PHYSICS,
UNIVERSITY OF ILLINOIS
June 20, 1924.

THE ELECTRIC MOMENT OF GASEOUS MOLECULES OF HALOGEN HYDRIDES

By C. T. Zahn

Abstract

Dielectric constant of the hydrogen halides HCl, HBr and HI to 300°C.—An improved heterodyne null method of measurement was used in which the beats between two high frequency oscillations from separate electron tube generators are adjusted to the frequency of a tuning fork, the small change of capacity due to the introduction of the gas being compensated by a large change in a large capacity in series with the gas condenser. Measurements were made at atmospheric pressure, using carefully purified gases, over a range of nearly 400° from just above the liquefaction point. The results are well represented by the Debye equation $(\epsilon-1)vT=AT+B$, where v is the specific volume and T the absolute temperature. [The values found for A are .001040 (HCl), .001212 (HBr) and .001856 (HI); and for B are .895(HCl), .52(HBr) and .12 (HI)]. This agreement lends support to the *Debye classical theory* of electric polarization due to fixed moments in the molecule. The values obtained for the *electric moment of the molecules* are, in 10^{-18} c.g.s.u., 1.034(HCl), 0.79(HBr) and 0.38(HI). The *Pauli quantum theory* gives an equation of the same form as Debye's but leads to smaller values for the electric moment. The upper limit for the moment given by infra-red absorption data for HCl, however, is 6 times the classical value and 13 times the quantum value and hence does not decide between the two theories.

Dielectric constant for hydrogen, oxygen, nitrogen and air.—The values obtained for $(\epsilon-1)\times10^6$ at 0°C and 760 mm pressure are 265, 518, 581 and 572 for H_2, O_2, N_2 and air, accurate probably to within 1/2 per cent.

Introduction

THE primary object of this investigation was to determine the electric moment of the molecules of HCl and HBr in order to compare their effective dipole length with their nuclear separation as obtained by Kratzer[1] from infra-red absorption spectra. This comparison would give, on certain assumptions, an idea of the distortion of the electronic configuration of the molecules. Further it was hoped that it might be possible by the use of high frequency technique to develop a method of measurement of dielectric constants more accurate than had hitherto been attainable by the use of electrometer and bridge methods and high frequency methods of different character. Similar methods have been

[1] Kratzer, Zeit. f. Phys. 3, 389 (1920)

used by Whiddington[2] for the measurement of small displacements, by Herweg[3] for studies of the effect of field strength on the dielectric constant, by Belz[4] for the measurement of the magnetic susceptibilities of salts, and by others. An experimental value of the dipole length also offered a possibility of choosing between the Debye theory and a more recent theory by Pauli applicable to dipole gases and based on Sommerfeld's conditions of quantum mechanics.

Debye's equation is as follows:[5]

$$\frac{\epsilon-1}{\epsilon+2}=\frac{4\pi}{3}N\left\{\lambda e^2+\frac{\mu^2}{3RT}\right\}=N\left\{A+\frac{B}{T}\right\},$$

where ϵ is the dielectric constant, N the number of molecules per cubic centimeter, λ the constant of elastic binding of the "optical" electrons, e the electronic charge, μ the electric moment of the molecule, R the Boltzmann gas constant and T the absolute temperature. The Debye equation has been found satisfactorily to represent data obtained by Jona and by Bädeker on H_2O, NH_3, SO_2, CH_3OH, C_2H_5OH, CS_2, CO_2 and CO.[6,7,8] Both Bädeker and Jona worked at temperatures in the range from room temperature to about 180° C. From the Debye equation it will be seen that the greatest contribution of the doublets, and so the greatest anomaly in the behavior of the dielectric constant is to be expected at the lowest temperatures. A more accurate verification of the Debye equation becomes of more particular interest also in view of a theory proposed more recently by Lundblad,[9] who considers the mutual effect of electrons of the same molecule.

Experimental Method

The measurement of the dielectric constant of gases is rendered extraordinarily difficult by the calibration of the apparatus. In fact, a number of investigators, some of whose results are quoted in tables of constants, have relied upon previous measurements for this calibration. Relative values of $(\epsilon-1)$ can be obtained much more accurately than absolute values. Indeed, different investigators have differed by as much as ten per cent in the values of gases such as nitrogen, air and hydrogen, but

[2] R Whiddington, Phil. Mag. **40**, 634 (1920)
[3] J Herweg, Zeit. f. Phys. **3**, 36 (1920)
[4] V. H Belz, Phil. Mag. **44**, 479 (1922)
[5] P. Debye, Phys. Zeit. **13**, 97 (1912)
[6] K. Bädeker, Zeit. f. Phys. Chem. **36**, 305 (1901)
[7] M. Jona, Phys. Zeit. **20**, 14 (1919)
[8] J J. Thomson, Phil. Mag. **27**, 757 (1914)
[9] R. Lundblad, Zeit. f. Phys. **5**, 349 (1921).

since the values for the different gases obtained by two observers usually bear a nearly constant ratio, the variations in the absolute values can be attributed to errors in the calibration of the apparatus. It is believed that the method herein described has reduced this difficulty of calibration of the apparatus and has accordingly given results less liable to such errors.

(a) *The measuring circuits.* The method of measurement adopted involves the use of high frequency oscillations produced by electron tube generators. A diagram of the electric circuits is shown in Fig. 1. There are two separate generators; one operates steadily at a frequency of the

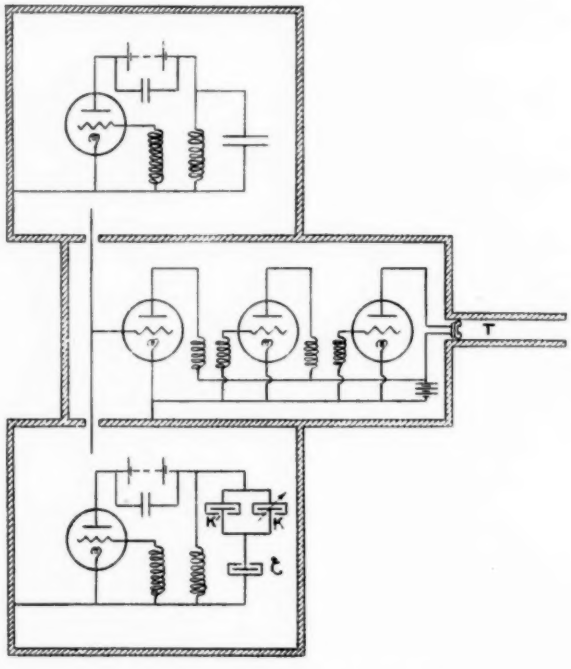

Fig. 1. Diagram of electric circuits.

order of a million per second; the other, at a frequency differing from this by an audible frequency of a thousand per second. These two generators are coupled electrostatically to a detector-amplifier circuit consisting of three stages as shown. A heterodyne note of frequency one thousand per second is heard from the telephone receiver *T*. This note is kept adjusted by the method of beats with a tuning fork operated by an auxiliary electron tube circuit. The variable generator contains a system of condensers one of which can be evacuated or filled with the gas to be

studied. The procedure is as follows. With the gas condenser evacuated the variable generator is adjusted to the condition of zero beats between the notes from the tuning fork and the receiver T; then the gas is introduced into the condenser and a measurable compensation is made for the resulting change in capacity by means of a device to be explained later.

The various parts of the electrical apparatus are carefully shielded wherever necessary. The two generators and the detector-amplifier are separately enclosed in metal boxes. In order to avoid coupling between the two generators themselves, and thus to make the frequency of the generators independent of each other, the coupling to the amplifier is reduced to a minimum. It is found quite sufficient to use a "floating grid"; i. e. a wire connected to the grid of the detector tube and having two branches, one passing through a small hole in the shield of each generator. This wire receives sufficient charge, mainly by electric induction, to give a conveniently audible note in the receiver T. The receiver itself is completely shielded electrostatically, the sound being heard at the end of a metal tube enclosing it. All the variable condensers are operated by shafts running through the shielding boxes, and readings are made through small holes in the boxes.

All five electron tubes are Western Electric 203B tubes operated at 45 volts on the plate and about 3 volts on the filament. The four inductance coils of the two generators consist of about 20 turns of 10 cm radius. For coupling between the stages of the amplifier, ordinary telephone transformers are used. The plate batteries of the generators are shunted with telephone condensers in order to reduce the resistance of the circuits for high frequency currents. All connecting wires are made of solid brass or copper rod.

(b) *The compensation device.* The system of condensers used in the variable generator consists of the three condensers K, K', and C shown in Fig. 1. K' is a fixed condenser of about $10000\mu\mu$f and K is a variable precision condenser of about $1500\mu\mu$f maximum capacity. The precision condenser has a slow adjustment screw which permits readings of capacity to be taken to within $(1/20)\mu\mu$f. The capacity $(K+K')$ is connected in series with the gas condenser C. Compensations for changes in C by the variation of K can be measured accurately since it will be seen that they can be made as large as desired according to the value of the shunted capacity K'. It would, of course, be quite impossible to measure the small capacity changes due to the influx of the gas into the condenser by means of an ordinary condenser in parallel with C. The series arrangement of the condensers magnifies the effect to be measured. Since by the

use of zero beats as a criterion of adjustment the total capacity of this combination is always kept constant it follows that for a compensation ΔK corresponding to a change ΔC produced by the influx of gas

$$\frac{1}{K+K'}+\frac{1}{C}=\frac{1}{K+K'+\Delta K}+\frac{1}{C+\Delta C},$$

or

$$\Delta C=-\left\{\frac{C^2}{(K+K'+C)\Delta K+(K+K')^2}\right\}\Delta K.$$

However, for the purpose of simplifying the calculations, a much more convenient form can be obtained for ΔC by the use of an approximate equation, which gives the same result as the latter equation to an accuracy within the limits of experimental error. For very small changes one could use the differential equation $dK/dC=d(K+K')/dC=-(K+K'/C)^2$, which, however, is not a sufficiently close approximation for the requirements of this problem. The next order of approximation consists in assuming that the differential coefficient dK/dC is a linear function of K. Then an average value of dK/dC can be used and for finite values of ΔC and ΔK

$$\Delta C=-\left\{\frac{C}{K+\tfrac{1}{2}\Delta K+K'}\right\}^2\Delta K.$$

For cases where measurements were to extend over a long period of time a special arrangement was used to eliminate errors due to gradual changes in the tube circuits external to the condenser system. This consists in having a standard condenser so arranged that it can be interchanged for the condenser system shown in Fig. 1. The interchange is made by a switch of special construction consisting of three equidistant mercury cups. Into one of these is inserted a vertical wire free to turn about the vertical axis and having attached a hozirontal wire which makes contact in turn with the mercury meniscus of either of the other two cups In order to compensate for possible changes in the circuits with time the distributed capacity of one of the generator circuits is varied by means of a threaded brass rod with a small brass disk attached to it on the inside of the shield. Small changes in frequency can be obtained by turning this rod and thus varying the distance of the small disk from the generator circuit.

(c) *The gas condenser.* The condenser C is a cylindrical condenser of about 4 cm diameter, 13 cm length and 0.5 mm separation between cylinders. Fig. 2 shows the main features. The two cylinders are held together and insulated by two quartz disks Q, which give a minimum of insulating dielectric to be corrected for in the calibration of the apparatus.

In order to prevent chemical action the cylinders themselves were made of an alloy of gold with small quantities of palladium and platinum. A gold plated brass condenser was first tried, but the distillation of zinc contaminated the gold plate at the higher temperatures. Around the outside cylinder is wound a platinum resistance thermometer P, insulated by a sheet of mica M wrapped around the condenser. The condenser is mounted in a Pyrex tube and held in position by tungsten wires fastened at the top and bottom of the gold cylinder. (These are not shown.) The two gold leads L and the two platinum leads P are first welded to tungsten wires T and then the seals in the glass tube are made at the top with the tungsten wires. A capillary tube is led from the tube to the vacuum and gas system. The Pyrex tube is then mounted in a small shielding cylinder of brass (plated with nickel to prevent corrosion by liquid air). This shielded condenser is then mounted in a large shielded box to permit the application of cooling baths and of the electric furnace without introducing capacity effects with the rest of the apparatus.

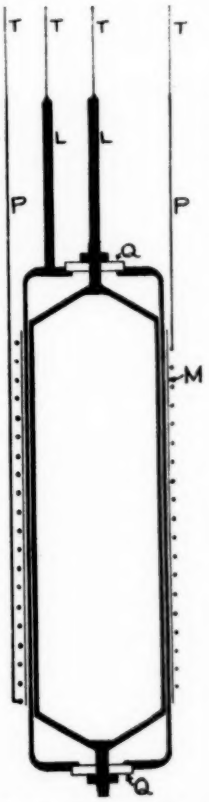

Fig. 2. Gas condenser.

The resistance thermometer consists of about 15 ohms of platinum wire wound around the condenser as loosely as is compatible with its staying in position. By means of a Wheatstone bridge the resistance can be measured to an accuracy of 0.001 ohm which corresponds to 0.02°C. The thermometer was calibrated under the experimental conditions to be used by means of three fixed points; the freezing and boiling points of water and the equilibrium point of a mixture of carbon dioxide snow and ether. In consideration of possible errors due to the calibration it is believed that this thermometer records temperatures accurate to at least 0.1 degree.

(d) *The calibration of the condenser system.* It will be noted that this is a null method and it depends in no way upon either the electron tube constants or the distributed capacity of the coils, etc., provided these do not change erratically. Actually, the limit of accuracy of the method depends upon the calibration of the condenser system. To make this as accurate as possible, the condenser system was so constructed that the whole calibration could be made in terms of one single condenser, the

precision condenser K. The latter was calibrated by the Bureau of Standards in terms of the standard air condensers of the Radio Section. The capacity K' consists of three fixed mica condensers arranged so that they can be inserted in parallel with K by means of mercury cup contacts. These capacities K' were such values that they could be calibrated in terms of K by the method of substitution. This has the advantage of preventing errors which might result from the dependence of the capacity of the fixed mica condensers upon the frequency. Here they are calibrated at precisely the frequency at which they are to be used. The capacity C is determined by shortcircuiting K and K' by amalgamated contacts and tuning the fixed generator to zero beats. Then C is shortcircuited and K is substituted and tuned thus giving a measure of C. With this arrangement an error of the unit in the calibration of K would produce no error in e so long as relative values of K are correct. To check the calibration of K a small fixed mica condenser was constructed and differences on K were checked with satisfactory agreement, so that settings on K are good to at least $\frac{1}{2}\mu\mu$f.

In the calibration of the condenser system the greatest possibility of error resides in the determination of the portion of the capacity C which is not altered by the influx of the gas. This is in part due to the quartz disks and in part to the capacity with ground of the lead passing from C to $(K+K')$. The inductance of this lead, even if appreciable, behaves just as though it were outside the system and in series with it and does not come into consideration. This lead will possibly have a small specific inductive capacity relative to the shield KK' and also with the grounded shield of C and other grounded apparatus such as the large shielding box enclosing the whole generator circuit. The former capacity is negligible as determined by both experiment and approximate calculation. The latter capacity is not negligible; it consists chiefly of the capacity of two wires about 15 cm long and 2 cm apart. These capacities were determined by a rather laborious process to be described below, by constructing a set of leads to imitate the actual leads to be used on the condenser tube and making measurements on these placed in position within the shields mentioned above. The lead capacity was thus found to be $5.5\mu\mu$f. To obtain the alterable capacity this must be subtracted from the total capacity $C=204.5\mu\mu$f, and in addition there must be subtracted the capacity due to the portion of the dielectric space occupied by the quartz disks. A calculation for this latter capacity gives $0.5\mu\mu$f. Hence the final value for the alterable capacity, which may be called $C_0=204.5-5.5-.5=198.5\mu\mu$f. The total correction for the lead and the quartz disk is less than three per cent of the total capacity C. A fair

estimate of the greatest possible error in this determination would be about $1\mu\mu f$ or 1/2 per cent.

The final equation for the dielectric constant is then:

$$\epsilon - 1 = \frac{\Delta C}{C_0} = \frac{C^2}{C_0} \frac{\Delta K}{\{K + \frac{1}{2}\Delta K + K'\}^2} = 210.6 \frac{\Delta K}{\{K + \frac{1}{2}\Delta K + K'\}^2}$$

The method of determining the lead capacity will now be considered in detail. The condensers K' were disconnected and condenser C was not yet mounted in its position in the shield S in Fig. 3a. L represents the immitation lead, at the end of which is connected by an amalgamated contact A another wire which passes out through the opening made by taking off the bottom of the shield S. The contact A is made or broken by a sidewise displacement of the lower wire. G is the ground capacity of the shield of K; and x and y, the lead capacities mentioned. C_1 is

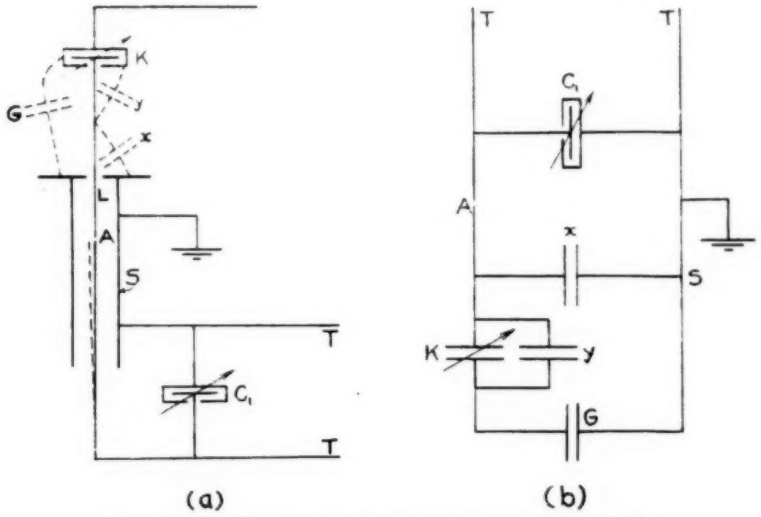

Fig. 3. Connections used in determining lead capacity.

another precision condenser calibrated in terms of K; and T, the terminals of the coil of a high frequency generator which produces beats with the fixed generator. In Fig. 3b this condenser group is redrawn symbolically for the sake of clarity. Now if the circuit is broken at A and a compensation $\Delta_1 C$ is made in C with the capacity K set at a particular value

$$\Delta_1 C = x + G(y+K)/(G+y+K)$$

where x, y, and G are unknown and $\Delta_1 C$ and K are obtained from condenser readings. By taking values of $\Delta_1 C$ for three settings of K there are obtained sufficient equations to solve for the three unknowns. A simpler and more accurate method is to use one equation of the latter type and

two other equations corresponding to the compensations $\Delta_2 C$ and $\Delta_3 C$ when G and $(K+y)$ are shortcircuited respectively. Then the following three much simpler equations are obtained

$$\Delta_1 C = x + G(K+y)/(G+K+y)$$
$$\Delta_2 C = x + K + y$$
$$\Delta_3 C = x + G.$$

The values of K, x, and y were obtained in this way and the method was checked by using different settings of K. It was in that way that y was found to be negligible.

(e) *The vacuum and gas system.* By the use of a diffusion pump the gas condenser could be evacuated within three minutes to a pressure as low as 10^{-3}mm Hg, which is far beyond the requirements of these experiments. The use of such a good pump is however important in ridding the system of water vapor, which is particularly necessary in the study of acids because of catalytic action. Before making a series of measurements the tube was first baked for several hours at a temperature of 300°C and evacuated to a pressure of 10^{-4} mm Hg or better. By taking this precaution and using P_2O_5 and a liquid air trap no difficulties were encountered such as prevented Bädeker from making a satisfactory study of HCl gas. Special arrangements were made for the generation and purification of each gas. These will be described in connection with the data on the individual gases.

Measurements on Air, Nitrogen, Hydrogen, and Oxygen

In order to test the behavior of the circuits, and to obtain data for comparison with those of other observers, measurements were made at room temperature on air, nitrogen, hydrogen, and oxygen. The air was passed through a P_2O_5 tube to absorb water vapor and through a liquid air trap to freeze out the CO_2. The nitrogen was prepared by heating sodium nitrite and ammonium chloride solutions together and freed from oxygen by storing it in a reservoir containing yellow phosphorous. It was then passed through a P_2O_5 tube with a liquid air trap. The same process of purification was used for hydrogen, which was prepared electrolytically. The oxygen also was prepared electrolytically and then passed through a P_2O_5 tube for drying.

A considerable number of readings was taken for these gases as shown in Table I. For all these individual readings the deviation from the mean is not greater than 1/2 per cent of the value of $(\epsilon-1)$ and the probable error of the observations is of the order of 1/10 per cent. The average value of the series capacity $(K+K')$ is 10000$\mu\mu$f which gives a magnification of 2500 in ΔC. The measured capacity ΔK is about 200$\mu\mu$f. All of

the measurements were taken at room temperature and approximately room pressure. Then the values were reduced to 0°C and 760 mm Hg by the use of the perfect gas law $pv = RT$. This latter procedure is justified by the fact that these gases are known to show no anomalous behavior of the dielectric constant.

TABLE I

Results for air, N_2, H_2 and O_2

Gas	No. of readings	$(\epsilon-1)$ reduced
Air	39	.000572
N_2	41	.000581
H_2	19	.000265
O_2	9	.000518

It is intended to give in the future a discussion of these data as related to the values of other investigators and to data on the refractive index in view of the Maxwellian relation. Suffice it to say here that these values agree with those of other investigators as well as the latter agree among themselves.[10] It is, however, believed that, in consideration of the improved method previously described, these values are reliable to within one per cent.

MEASUREMENTS ON HCl, HBr, AND HI

After the preliminary measurements had been completed a series of observations were made on the gases HCl, HBr, and HI. The first method of procedure tried with HCl was unsatisfactory. The gas was let in after cooling the apparatus almost to the liquefaction point of HCl. Then readings were taken on the condenser K as the apparatus approached room temperature. This required several hours and it was found after-

[10] In anticipation of later comment, it seems desirable to call attention to a recent paper by Fritts (Phys. Rev. **23**, 345, 1924) which gives data on the dielectric constants of oxygen, nitrogen and air differing by several per cent from those of the present paper, although claiming about the same accuracy. Fritts used a very ingenious method of recording the beats, so that his actual observations are certainly very accurate. His precautions regarding purity of gas and calibration of apparatus do not seem to be equally trustworthy, as judged by experience during the present investigation. We found the step-by-step calibration, in addition to the Bureau of Standards calibration, necessary in order to realize the possibilities of the method. We also found the method so sensitive as to detect impurities left in the gases after ordinary chemical methods of purification, so that unusual methods were employed to advantage. The correction for lead capacity is also important and difficult to be sure of. Since the present apparatus was designed primarily to permit convenient temperature adjustments, a new apparatus is being used to investigate further the sources of error in *absolute* values. For this reason further discussion of absolute values is deferred until a later communication.

ward that the gas had become contaminated, either as a result of chemical action or by gases occluded in the condenser tube. It was therefore found necessary to use another method of procedure, wherein after each gas reading a vacuum reading was taken and impurities in the gas were removed by a method described below. Formerly another source of error was found in the fact that whenever the temperature of the condenser was changing rapidly, the temperature of the inner cylinder would lag behind that of the outer, particularly when the condenser was evacuated, because of the extremely good heat insulation provided by the quartz disks; for this reason it was found necessary to take observations after steady temperatures had been reached or under conditions of slowly varying temperature. Further by the first method it was necessary to make an experimental run with the condenser evacuated in order to determine the change in K due to the thermal expansion of the gas condenser. By the second method of procedure the importance of this temperature effect is eliminated since now the effect appears only as a change in C and C_0, whereas before it appeared also as a correction in ΔK, a quantity of the same order of magnitude as the correction. In C there is produced thereby a change of the order of 1/10 per cent, which is neglected in the computations.

The preparation and purification of the three gases was achieved in the following manner. HCl was generated by dropping concentrated H_2SO_4 slowly onto concentrated HCl. The gas thus generated was passed consecutively through concentrated H_2SO_4, a P_2O_5 tube, and a liquid air trap where it was frozen. Then the liquid air trap was evacuated over the frozen HCl to remove any non-condensible gases such as air and hydrogen. Before letting the gas into the condenser tube a portion of it was boiled off and removed in order to eliminate possible impurities of lower boiling than that of HCl. Then the HCl was allowed to boil and to pass into the condenser until atmospheric pressure was reached, the stopcock to the generator was opened, and the surplus gas was allowed to bubble away through a trap of concentrated H_2SO_4. Other possible impurities would boil off last and have little chance of diffusing through the capillary tube running to the condenser. In this way the same gas could be used again and again. It may be noted that on renewing the gas no change in ϵ occurred.

For the generation of HBr bromine was dropped slowly into a moist mixture of sand and red phosphorous, and the gas evolved was passed through a tube of red phosphorous to remove free bromine vapor and to the P_2O_5 tube and the liquid air trap where it was distilled as was the HCl. The surplus gas was bubbled through concentrated HBr.

HI was generated similarly by dropping a thin mixture of water and red phosphorous onto moist iodine. It was bubbled through concentrated HI to remove free iodine vapor and then dried and frozen as was done for the other two gases.

Table II shows the experimental and calculated data for these three gases. Each value $(\epsilon-1)$ represents the mean of from 5 to 10 or more readings taken at the given temperature. In the third column is given the value $T/273$, which is the specific volume relative to 0°C and 760mm Hg and calculated according to the perfect gas law. The fourth column gives the correction to be applied to this value of v as obtained from the constants of the van der Waals equation; and the fifth column, the corrected value of v. In the sixth column is given the value of $(\epsilon-1)$ reduced by a slight correction for pressure made on the simple gas law. The next column shows $(\epsilon-1)v$, which is the value of $(\epsilon-1)$ reduced to constant density by use of the Clausius-Mossotti relation $(\epsilon-1)v=$const. for a given temperature. These values of $(\epsilon-1)$ for constant pressure and for constant density are shown in Fig. 4 to give an idea of the order of magnitude of the variations occurring throughout the range of temperature studied.

TABLE II

Gas	$T°$ K	$\frac{T}{273}$	Correction to v	v	$(\epsilon-1)_{760}$	$(\epsilon-1)v$	$(\epsilon-1)vT$
HCl	201.4	.7375	−.60%	.733	.007452	.005462	1.110
	260.1	.953	−.07	.9527	.004716	.004494	1.169
	294.2	1.078	+.08	1.079	.003792	.004092	1.205
	359.2	1.316	+.22	1.319	.002672	.003524	1.266
	433.9	1.590	+.23	1.594	.001948	.003104	1.347
	503.9	1.847	+.23	1.851	.001526	.002823	1.423
	588.8	2.157	+.27	2.162	.001182	.002555	1.505
HBr	217.5	.7968	+.45	.7925	.004545	.003602	.7835
	269.0	.985	−.01	.985	.003209	.003160	.8501
	294.4	1.079	+.08	1.080	.002796	.003020	.8840
	338.9	1.241	+.36	1.245	.002228	.002773	.9400
	478.5	1.753	+.54	1.762	.001303	.002295	1.098
	599.1	2.194	+.59	2.208	.000943	.002081	1.248
HI	244.5	.8955	*		.002687	.002402†	.5885
	295.1	1.081			.002123	.002296	.6776
	345.9	1.267			.001760	.002231	.7713
	474.6	1.738			.001239	.002153	1.022
	612.2	2.243			.0009706	.002177	1.334

* No values of the van der Waals constants are available for this correction.
† The uncorrected value of v is used here.

The values in the last column of Table II serve as a test of the Debye equation as follows. This equation can be written as

$$(\epsilon-1)vT=AT+B$$

which shows that if the values of the last column be plotted against the absolute temperature there should be obtained a straight line with intercept B determined by μ, the electric moment, and with slope A depending upon the contribution of the elastic electrons to the polarization.

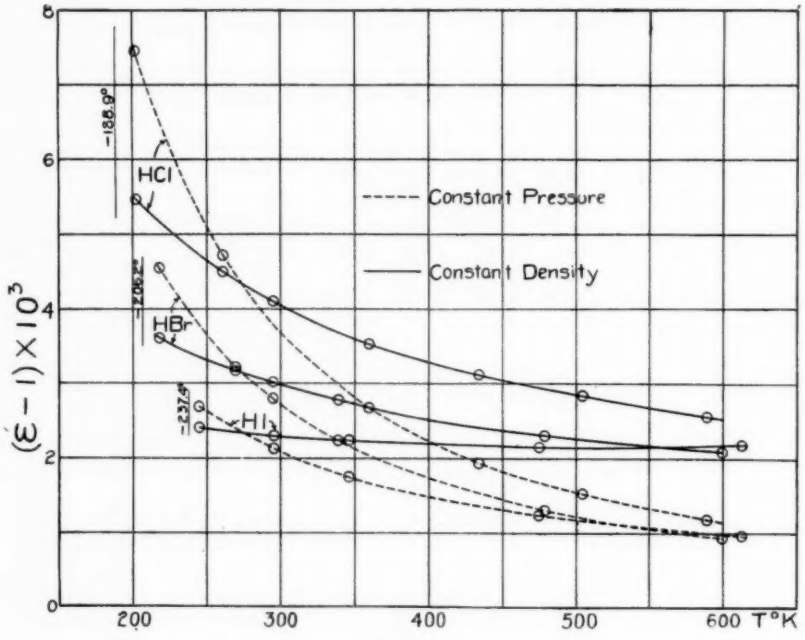

Fig. 4

These points are plotted in Fig. 5. For HCl and HBr straight lines could be drawn so that none of the points deviate by more than 1/2 per cent of $(\epsilon - 1)$, even at temperatures near the liquefaction point. This agreement cannot be shown adequately by the necessarily small scale drawing of Fig. 5. For HI two sources of difficulty arise. At high temperatures this gas dissociates to a marked degree. In Gmelin and Kraut[11] the equilibrium degree of dissociation of HI is given as about 3 per cent at 0°C and 18 per cent at 300°C. In the experiments described this degree of dissociation was probably not reached at the moderately high temperatures on account of the slowness of the reaction velocity at such temperatures. (Any dissociated gas was removed after each reading by the purifying process so that no accumulation of dissociated gases could take place.) Since the number of molecules does not change in the dissociation, except possibly by a slight amount due to condensa-

[11] Gmelin und Kraut, Handbuch der anorganischen Chemie **1**, (2), p. 321.

tion of the iodine on the walls of the tube, this consideration becomes less important. Also from data on the refraction indices of HI, H_2, and I it is seen that the refractivity of HI is about ten per cent greater than would be expected on the principle of additivity of refractivities. Hence in dissociation the effect of the destruction of the moment would in part be compensated for by the increase in the contribution from the elastic electrons. However, the experimental values of $(\epsilon-1)$ give a much better straight line than might have been hoped for. With the above considerations in view a straight line was drawn through the lower three points,

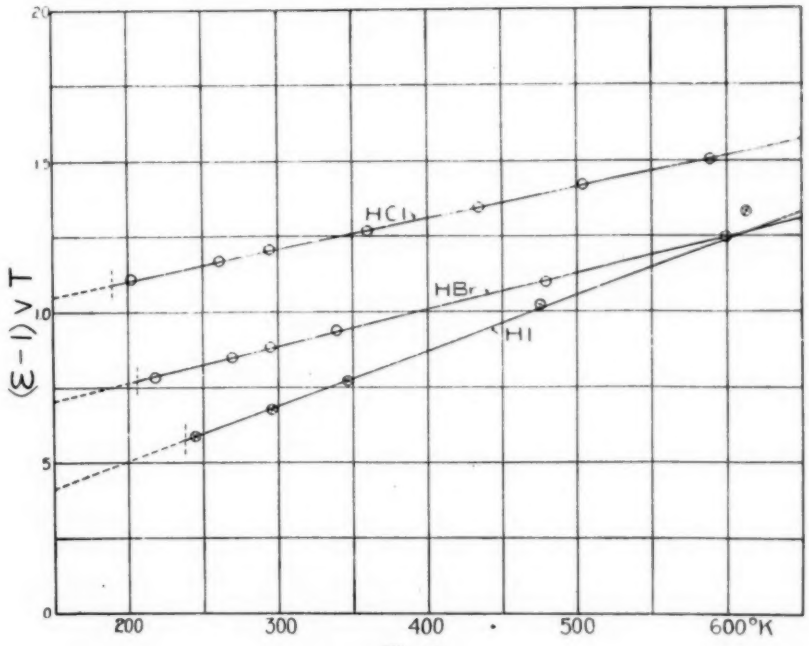

Fig. 5

where very probably the above mentioned errors are only very small; the fourth point is fairly good but the last point is considerably off the straight line. Another difficulty in the case of HI is that no data are available on the constants of the van der Waals equation. The uncorrected values of v were used here. This introduces a possible error of about 1/2 per cent.

Table III gives the results obtained from the lines of Fig. 5. The second column under $(\epsilon-1)_0$ gives the contribution of the elastic electrons to $(\epsilon-1)$; and the third column, values of (n^2-1) for the Na D line, since (n^2-1) for long wave-lengths is not known. These two columns should

be identical if the values of the refractive index were given for long wavelength instead of for the Na D line. The last column gives the values of the electric moment as calculated from the intercept of the lines of Fig. 5.

$$B=(4\pi/3)(N_0\mu^2/3R), \quad \text{(using } \epsilon+2=3\text{)}$$

or $$\mu^2=(3\times1.36\times10^{-16}/4\pi\times2.71\times10^{19})B=1.198\times10^{-36}\,B$$

TABLE III

Results for HCl, HBr and HI_2

Gas	$(\epsilon-1)_0$	(n^2-1)	$\mu\times10^{18}$
HCl	.001040	.000888	1.034
HBr	.001212	.001140	.788
HI	.001856	.001812	.382

DISCUSSION OF DATA

Of the three dipole gases here studied only HCl has been worked upon before so far as is known by the author. Bädeker[6] attempted to study HCl over a considerable range of temperature but was prevented by chemical action which he attributed to moisture. However, his values of $(\epsilon-1)$ for 84.4° and 106.6°C, .000269 and .000256 show very good agreement with the corresponding values from Fig. 4, .000270 and .000247, better than might be expected considering the fact that Bädeker did not put much confidence in his values. These values for HCl do not extend over a sufficiently large temperature range to allow the calculation of the electric moment with any degree of certainty.

It should be mentioned here that recently Falkenhagen[12] has published a value of the electric moment of HCl, $\mu=2.15\times10^{-18}$ c.g.s.u., obtained from dielectric constant measurements. This is about twice the value given in Table III. Since no values of ϵ are given and only a brief description of the method, it is impossible to arrive at an explanation of the discrepancy. A calculation has been made of the contribution of the doublets alone to $(\epsilon-1)$ using the value given by Falkenhagen for the moment. This gives $B=\mu^2\times10^{36}/1.198=3.86$; hence $B/T=3.86/373=.0104$ for 100°C. Thus it is seen that Falkenhagen's value of the moment leads to a value of electric polarization about four times the total value observed by the author and also by Bädeker. Hence the value given by Falkenhagen would seem to be impossibly large.

Frivold and Hassel[13] have made a determination of the dipole length of the HCl molecule by means of electrostriction. They give for the value of the dipole length 0.31×10^{-8} cm. This corresponds to a value of

[12] H. Falkenhagen, Phys. Zeit. **23**, 87 (1922).

[13] O. E. Frivold und Odd Hassel, Phys. Zeit. **24**, 82 (1923)

$\mu = 1.48 \times 10^{-18}$ c.g.s.u., which is considerably less than the value of Falkenhagen. These electrostriction measurements are of very limited accuracy.

In support of the value obtained by this investigation Smyth[14] has calculated the electric moment of HCl by a simplified use [15] of an equation developed by Gans.[16] This equation applies to both the gaseous and the liquid states and is developed on the dipole assumption taking into account the intermolecular forces. By Smyth's method the electric moment can be calculated approximately from a knowledge of the indices of refraction at two different wave-lengths, the corresponding density, the dielectric constant at one temperature, the density at this temperature, the molecular diameter, and certain universal constants. For the value of the dielectric constant he used Bädeker's two values for the gaseous state and Schlundt's value for the liquid state at 27.7° C and obtained for $\mu \times 10^{18}$ at these three values, 1.096, 1.054, and 1.075 c.g.s.u. respectively. The mean of these three values, 1.075, agrees with the value of Table III, 1.034, better than could be hoped for taking into consideration possible errors in the calculations and in the experimental values involved.

A further argument in favor of the value 1.034 is the fact that $(\epsilon - 1)_0$ in Table III agrees fairly closely with the refractive index data.

Relation to Infra-Red Spectra

Values of the electric moment of the halogen hydrides are of particular interest for two reasons; first their structure is perhaps the simplest dipole structure because of the simple hydrogen atoms, and second their nuclear separation (except for HI) has been determined from infra-red absorption data.[17] Now it will be noted that once the nuclear separation is known a knowledge of the average position of the electronic configuration obtained from the value of the electric moment would be of considerable importance in the study of the structure of the halogen hydrides. Suppose that in the formation of the HCl molecule, for example, the electron from the hydrogen atom combines with the seven outer electrons of the chlorine atom to form a group similar to the outer group of the argon atom. If this argon group be assumed to be symmetrical and not appreciably distorted by the presence of the hydrogen nucleus, then the

[14] This work is not yet published.
[15] C. P. Smyth, Phil. Mag. **45**, 850 (1923).
[16] R. Gans, Ann. der Phys. (4) **44**, 481 (1921);
H. Isnardi and R. Gans, Phys. Zeit. **22**, 230 (1921).
[17] See A. Kratzer, Zeit. f. Phys., **3**, 305 (1920);
Herbert Bell, Phil. Mag. p. 549, March 1924

chlorine atom plus its acquired electron would be equivalent to a single negative charge at the chlorine nucleus. This latter charge combined with the hydrogen nucleus gives an effective moment of HCl such that its dipole-length is equal to its nuclear separation. But one would expect the hydrogen nucleus to distort the electronic orbits in such a way as to decrease the dipole-length. Hence it seems reasonable to regard the nuclear separation as an upper limit to the dipole-length as pointed out by Pauli. C. P. Smyth[18] has made a calculation for the effect on the electric moment of the HCl molecule due to the distortion by the hydrogen nucleus. For this purpose a statical structure of the Langmuir type was assumed and the distortion of the individual electrons was calculated from the coefficient of elastic binding given by the refractive index of the substance. These calculations give the proper order of magnitude for this distortion. From the point of view of the Bohr theory one would not expect a calculation based upon the above assumptions to give a better agreement.

Such a consideration is of interest in connection with the quantum theory of dielectric polarization developed recently by Pauli.[19] Whereas Debye and Thomson used in the development of their formulas classical mechanics, Pauli has derived for the case of dipole gases a formula based on Sommerfeld's quantum conditions. The latter theory leads to an equation of the same form as that of Debye, differing only in the numerical factor which appears in the doublet term after integration.

$$\frac{\epsilon-1}{\epsilon+2}=(4\pi/3)\ N\{\lambda c^2+1.5367\ \mu^2/RT\}\ .$$

Thus so far as moments are concerned the two theories would give different values for the electric moment calculated from data on the dielectric constant, such that:

$$\mu\ (\text{classical})=2.15\ \mu\ (\text{quantum}).$$

The measurement of the dielectric constant of dipole gases at different temperatures offered a possibility of distinguishing between the two theories, should the classical theory lead to a value of the dipole-length greater than the nuclear separation.

In Table IV are given the results of some calculations relating to the above mentioned quantities. d is the dipole-length calculated on the classical theory, and r_0 is the nuclear separation obtained from the infra-red data. D is the distance of the center of gravity of the electronic configura-

[18] C. P. Smyth, Phil. Mag., **47**, March 1924
[19] W. Pauli, Jr., Zeits. f. Phys. **4**, 319 (1921)

tion from the halogen nucleus; this may be regarded as a measure of the distortion caused by the hydrogen nucleus. A simple calculation gives $D=(r_0-d)/(N+1)$, where N is the atomic number of the halogen atom. Since the value of the dipole-length is about one-sixth the value of the nuclear separation it is impossible to choose between the quantum theory and the classical theory.

TABLE IV

Comparison of calculated dipole length d with nuclear separation r_0 from infra-red data.

Gas	$\mu\times10^{18}$	$d=\frac{\mu}{e}$	$r_0\times10^8$	$D=\frac{r_0-d}{N+1}$
HCl	1.034	.217	1.27	.0618
HBr	.788	.165	1.37	.0343
HI	.382	.080	....	.....

In conclusion the writer wishes to acknowledge a grant from the Joseph Henry Fund of the National Academy of Sciences, which has in part made possible the construction of special apparatus for this problem. The writer is also deeply indebted to Professor K. T. Compton whose stimulating interest and whose most generous help and advice have contributed largely to the success of this investigation.

PALMER PHYSICAL LABORATORY,
PRINCETON UNIVERSITY,
May 8 1924

THE MAGNETIC SUSCEPTIBILITY OF HELIUM, NEON, ARGON, AND NITROGEN

By L. G. Hector

Abstract

The apparatus described in a recent paper by Wills and Hector[1] for the measurement of the magnetic susceptibilities of gases, has been rebuilt so as to enable the temperature to be controlled to .02° and the sensitivity of the method increased. The gases used were all carefully purified, and spectroscopic examination showed no oxygen lines. The results obtained for the volume susceptibilities at 20°C and 760 mm pressure for helium, neon, argon and nitrogen are -0.780, -2.77, -7.52, and -4.91, respectively, all times 10^{-10}. They are probably accurate to within 2 per cent. These results are of the order of magnitude required by Pauli's modification of the Langevin theory of diamagnetism and also agree with values predicted by Joos from the electron arrangements.

Introduction

In a previous paper by A. P. Wills[1] and the writer a method was developed for determining the susceptibility of fairly strong magnetic gases by balancing them magnetically against a water solution of nickel chloride. It was further shown that the method could be extended to the case of the nearly neutral diamagnetic gases by balancing the gases magnetically at various densities (pressures) against the same solution at slightly different temperatures, since the magnetic temperature coefficient of the solutions was large and that of these gases relatively negligible. The method was carried out for the cases of hydrogen and helium, but only tentative results were obtained for the latter gas because of its extremely small value. The present research involves perfecting the details of the method in order to be able to determine such small susceptibilities as that of helium, and applying the method to the measurement of the other gases, argon, neon, and nitrogen.

It was shown in the previous paper that the working formula for this method of determining the susceptibilities of diamagnetic gases is:

$$\kappa_0 = -\frac{P_0\theta}{\theta_0}\,\frac{\rho_w \chi_w}{1+R-K}\,\frac{dR_0}{d\theta}\,\frac{\Delta\theta}{\Delta P} \qquad (1)$$

where κ_0 is the volume susceptibility at the absolute temperature θ_0 and standard pressure P_0 in mm mercury, and R_0 is the ratio of mass of

[1] A. P. Wills and L. G. Hector, Phys. Rev. **23**, 209 (Feb. 1924)

water to mass of nickel chloride for a magnetically neutral solution at the absolute temperature θ at which the experiment is performed. ρ_w is the density and χ_w the volume susceptibility of water at the temperature θ.

R is the actual ratio of the mass of water to mass of nickel chloride in the solution used, and may be replaced by R_0 without appreciable error for these nearly neutral solutions. K is a constant depending solely on the chemical concentration and density of the mother solution from which all the solutions used were made by dilution by weighing.

The mother solution used was the same as that employed in the previous research, and hence we have the same value of K, 0.88102. We have also the same Curie constant ($C = R_0\theta$) for this solution of nickel chloride. As before specified, there was about 0.08 per cent cobalt chloride in this solution, and the Curie constant obtained was C = 13890. The value of χ_w is taken to be -0.720×10^{-6}. If we insert these values in Eq. (1) we have for standard pressure (760 mm Hg) and a temperature of 20°C

$$\kappa_0 = -0.02594 \frac{\rho_w}{0.1190\ \theta + 13890} \frac{\Delta\theta}{\Delta P} \qquad (2)$$

where $\Delta\theta/\Delta P$ is to be determined by experiment and expressed in degrees per mm Hg pressure. The temperature θ must be observed at the same time and the corresponding value of ρ_w taken from a table of water density.

The method employed was to determine a fairly large number of values of θ and the corresponding values of P for a given solution. These values were then plotted, and the quantity $\Delta\theta/\Delta P$ determined from the graph by an approximate least squares solution.

Procedure

Oxygen coming from the air dissolved in the solution formed a principal source of trouble at the beginning of our attempts to measure the diamagnetic gases. However, by means of a vacuum pump a pressure of about 1 cm of mercury was held on the balance for about one half hour at the beginning of a run. This process hastened the removal of dissolved air from the solution. The apparatus was then flushed with hydrogen for from two to three hours, and then with the gas to be measured for about one half hour.

The apparatus was similar to that used in the previous research with the addition of a double-walled house about the entire magnet. The temperature of the air within the house was controlled to within plus or

minus 0.02°C by means of a bi-metallic strip which applied potential to the grid of a three electrode tube, in the plate circuit of which was a high resistance relay. Temperature adjustment was effected by means of a bank of lamps controlled by the relay, and by means of pails filled with ice and placed in openings in the top of the house. An electric fan kept all the air in the house in circulation.

The gases (except neon) were supplied to us under high pressure in steel cylinders. Pressures, except for neon, were read by means of an ordinary gauge calibrated against the mercury manometer described in the previous paper. Absolute values of temperature were read from a mercury thermometer calibrated to 0.02°C, placed between the poles of the magnet, very near to the balance. The small variations in the temperature were read from the constant deflections produced on a galvanometer by means of a copper-constantin thermopile. One set of junctions was kept in a Dewar bottle filled with ice and water, and the other set was wrapped about the horizontal tube of the balance and protected from the air by cotton wool and tape, so that variations in temperature would come from the liquid within. The temperature variations in the liquid were produced by slightly varying the temperature of all the air inside the double-walled house. These variations were small, seldom exceeding 0.2 or 0.3° at a time.

All the gases measured were examined spectroscopically for oxygen, but in no case were any oxygen lines observed.

Helium

The helium supplied to us contained about 8 per cent nitrogen. It was purified by passing it through a long tube filled with magnesium granules and kept at a dull red heat by means of an electric furnace. This procedure served to remove any traces of oxygen as well as the nitrogen.

The following tables give the data from which the graphs and calculations were made. Galvanometer readings and gauge pressure readings are omitted and only the corresponding columns of corrected pressures and temperatures are given. The numbers in the temperature column are in centigrade degrees, but are referred to an arbitrary zero which is different in the different tables. Pressures are in mm Hg at 0°C.

Fig. 1 shows the $(\theta - P)$ graph for only one set of data for each gas. All the values of $\Delta\theta/\Delta P$ given have, however, been taken from similar graphs of the various sets of data given in this paper.

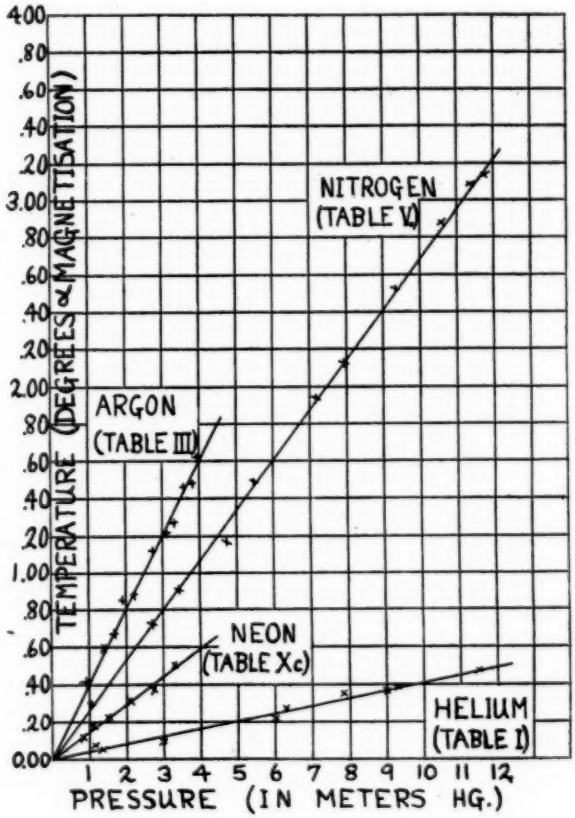

Fig. 1.

TABLE I

Results for helium

Pressure	Temperature
1230mm	.94°
6000	1.11
9000	1.26
11440	1.36
9280	1.27
7870	1.25
6200	1.15
2930	.99
1180	.96

TABLE II

Results for helium

Pressure	Temperature
1180mm	3.17°
4950	3.31
8450	3.46
10880	3.70
10875	3.64
10570	3.58
7770	3.51
5950	3.46
5200	3.43
3820	3.34
3220	3.32
2400	3.31
1230	3.26
930	3.25

From the data of Tables I and II with the aid of Eq. (2) we find the following values

	θ	ρ_w	$\Delta\theta/\Delta P$	κ_0
(I)	292.2°K	.9985	4.12×10^{-5}	-0.766×10^{-10}
(II)	294.1°	.9980	4.28	-0.795
			Mean	-0.780×10^{-10}

Argon

The argon as supplied to us contained about 4.5 percent nitrogen and 0.5 per cent oxygen. The purifying train consisted of heated tubes, one filled with magnesium granules, the other with copper filings which had previously been oxidized and reduced.

Table III

Results for argon

Pressure	Temperature
1680mm	1.86°
2250	2.11
3320	2.47
3875	2.69
3950	2.82
3580	2.67
3080	2.41
2750	2.34
1980	2.07
1380	1.80
990	1.62

Table IV

Results for argon

Pressure	Temperature
1375mm	.83°
4740	2.13
4920	2.38
4590	2.12
4180	2.03
3500	1.99
3120	1.79
2850	1.63
2600	1.46
2140	1.26
1680	1.07
1375	.91
1200	.86
1080	.74
900	.72

From the data of Tables III and IV, with the aid of Eq. (2), we obtain the following values

	θ	ρ_w	$\Delta\theta/\Delta P$	κ_0
(III)	291.7°K	.9985	4.10×10^{-1}	-7.63×10^{-10}
(IV)	290.1°	.9989	3.98	-7.41
			Mean	-7.52×10^{-10}

Nitrogen

The nitrogen as supplied to us was supposed to be better than 99.9 percent pure, but as a final precaution against oxygen we passed it through a heated tube filled with oxidized and reduced copper filings.

Table V

Results for nitrogen

Pressure	Temperature
4720mm	2.68°
7140	3.43
10570	4.38
11340	4.59
11700	4.63
9330	4.02
7970	3.64
5400	2.98
2700	2.23
1030	1.80
3380	2.40

Table VI

Results for nitrogen

Pressure	Temperature
3710mm	1.58°
6810	2.15
8220	2.63
9890	3.16
11340	3.48
11640	3.56
10410	3.28
8890	2.93
7250	2.48
5750	2.02
3250	1.45
1860	1.11
1230	.88

Table VII

Results for nitrogen

Pressure	Temperature
7250	3.44°
9950	4.33
11340	4.75
8030	3.93
5350	3.20
2700	2.49
1320	2.08

From the data of Tables V, VI, and VII, with the aid of Eq. (2) we find the following values

	θ	ρ_w	$\Delta\theta/\Delta P$	κ_0
(V)	291.2°K	.9986	2.68×10^{-4}	-4.98×10^{-10}
(VI)	290.6°	.9987	2.57	−4.78
(VII)	291.4°	.9986	2.67	−4.97
			Mean	-4.91×10^{-10}.

NEON

The neon was purified by passing it through a trap packed with charcoal and immersed in liquid air. Quantitative chemical analysis showed no oxygen present, but there was a small amount of residual nitrogen and some helium. These impurities were estimated at 2 per cent or less nitrogen and 3 per cent helium. Such a mixture would have a susceptibility so near that obtained for neon that any errors arising from considering the mixture to be 100 per cent neon would be within the experimental limits of precision. As a final precaution against oxygen a tube of hot reduced copper filings was kept in the gas train throughout the measurements.

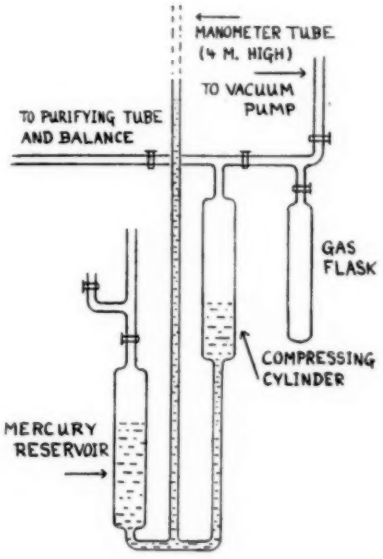

Fig. 2. Apparatus used in compressing neon.

The neon was supplied to us in 350 cc glass flasks at atmospheric pressure. Fig. 2 shows the method of compressing this gas in the balance without any opportunity of introducing impurities.

TABLE VIII

Results for neon

(a) Pressure	Temp.	(b) Pressure	Temp.	(c) Pressure	Temp.
2017mm	2.53°	1867mm	2.65°	3532mm	2.95°
1349	2.44	1290	2.62	2697	2.84
910	2.33	896	2.57	1808	2.75
1379	2.47	2236	2.72	871	2.60
1967	2.52	2824	2.82		
3572	2.82	3074	2.90		
3213	2.76	3562	2.95		
2655	2.72				
2117	2.63				

TABLE IX

Results for neon

(a) Pressure	Temp.	(b) Pressure	Temp.
808mm	1.77°	848mm	1.91°
908	1.80	1526	2.02
1906	1.91	2305	2.14
2964	2.11	3053	2.26
3313	2.19	3323	2.29
2405	2.06	2305	2.16
1576	1.95	987	1.95
848	1.84		

TABLE X

Results for neon

(a) Pressure	Temp.	(b) Pressure	Temp.	(c) Pressure	Temp.
836mm	1.86°	796mm	1.99°	816mm	2.03°
1892	1.99	1922	2.12	1484	2.13
2919	2.13	3248	2.34	2760	2.29
3288	2.27	3059	2.30	3284	2.41
2092	2.06	2172	2.19	2042	2.22
796	1.89	816	2.02	1055	2.09

From these data with the aid of Eq. (2) we find

		θ	ρ_w	$\Delta\theta/\Delta P$	κ_0
VIII	(a)	291.0°K	.9987	1.72×10^{-4}	-3.20×10^{-10}
	(b)	291.3°	.9986	1.56	−2.90
	(c)	291.4°	.9986	1.26	−2.34
IX	(a)	292.3°	.9985	1.51	−2.81
	(b)	292.4°	.9984	1.53	−2.85
X	(a)	292.8°	.9983	1.48	−2.75
	(b)	293.0°	.9983	1.35	−2.51
	(c)	293.0°	.9983	1.47	−2.73
				Weighted mean	$-277.\times10^{-11}$

In obtaining the mean, the mean values from each of the three tables were given equal weight.

TABLE XI

Volume susceptibilities at 20° C and 760mm pressure

	$\kappa_0\times10^{10}$
Helium	−0.780
Neon	−2.77
Argon	−7.52
Nitrogen	−4.91

The neon, argon, and nitrogen were made by fractional reduction of liquid air by the Air Reduction Company of New York City. I am indebted to Dr. F. J. Metzger of this company for providing me with exceptionally pure samples for our measurements. In the case of argon and nitrogen, which were furnished us under high pressure, water lubricated compressors were used to avoid contamination with oil vapors. The helium was provided by the Bureau of Aeronautics, Navy Department.

Sensitivity of Apparatus and Precision of Measurements

A variation of plus or minus about 0.03° from the ideal zero setting corresponded to the smallest magnetic difference for which a lack of magnetic balance could be observed. The corresponding change in the susceptibility of the solution is about 0.7×10^{-10}. With the field used (25,000 gauss) and a tube of diameter 4 mm, this gives a force of about 1/300 dyne as the limit of the sensitivity of the device. These figures give some idea of the dependability of the individual points determined. But the slope of a line on the (θ,P) graph is, of course, determined from a number of points. Judging from the agreement of the independent determinations, the mean values are probably accurate within ±2 per cent. The accuracy of the results depends also on the value of the susceptibility of water, which I have assumed to be -0.72×10^{-6} plus or minus one half of one per cent.

Discussion of Results

Pauli[2] has shown, by a modification of the Langevin theory of diamagnetism, that the order of magnitude of the susceptibility of helium and argon as based on the knowledge of atomic magnitudes from other sources should be less than -2×10^{-10} for helium at 0°C and between -2.7 and -8.1×10^{-10} for argon at 0°C. By a comparison of data on the susceptibilities of ions having similar electron arrangements, Joos[3] has also predicted values for these gases, and obtained -1.3×10^{-10} for helium at 0°C and -7.8×10^{-10} for argon at 0°C. An extension of his method and data would give -3.9×10^{-10} for neon at 0°C. The agreement between the experimental values and these predicted orders of magnitude is striking.

The writer is indebted to Professor A. P. Wills for advice and suggestions throughout the research.

Department of Physics,
Columbia University,
June 10, 1924.

[2] W. Pauli, Jr. Zeit. f. Physik **2**, 201 (1920)
[3] George Joos, Zeit. f. Physik **19**, 347 (1923)

CRYSTAL STRUCTURE OF HARD STEEL

By K. Heindlhofer

Abstract

Using an x-ray spectrometer with ionization chamber and reflecting Mo $K\alpha$ and $K\beta$ rays from small blocks of steel, spectra were determined for 0.80 and 1.31 per cent carbon steels both in coarse and in fine martensitic condition. The same lines were found as in the spectrum of carbon free iron, but they were only one-third as intense and were somewhat broader and slightly shifted. *Lattice of martensite.* These results confirm the previously reported conclusion that martensite has a body centered lattice like that of α iron but show that the lattice is contracted about 0.4 per cent on a side and is also distorted. This distortion, which changes the shape of the lines, disappears largely however after heating the martensite at 260°C for 30 min., although the hardness is not much decreased. *Nature of martensite.* From a comparison of the densities of the specimens with those of α iron and of cementite, it is concluded that martensite is a mixture (1) of a solid solution of carbon in iron, the C atoms replacing the Fe atoms, and (2) of finely dispersed cementite. The solid solution is unstable even at room temperature. Both constituents seem to contribute to the *hardness*, the solid solution by distorting the lattice and the cementite by being so fine grained.

Theories of the Hardness of Metals

THE knowledge of the mechanism of hardening of steel after being quenched from a certain definite temperature range is both of theoretical and practical importance. We shall recapitulate two important theories which seem best to explain hardness of alloys in general, very likely including steel as a special case.

(a) *The hardness of solid solutions explained by Rosenhain.*[1] It is a well known fact that metals which are mutually soluble in any proportions do not form alloys of great hardness as compared with that of the component metals. Rosenhain points out that the crystalline habit of such component metals is very similar. For instance gold and silver are soluble in any proportion, their alloys are soft, and both crystallize in face centered cube of almost identical size. In this and similar cases the atoms of the one metal indiscriminately replace the atoms of the other, without materially disturbing the space lattice.

There are, however, a great number of metals, which show a limited solubility. This is because when some of the atoms of the metal A are replaced by the atoms of the metal B, the lattice is disturbed due to

[1] Walter Rosenhain, Proc. Roy. So . A **99**, 196 (June 1921)

disturbed symmetry. Distortion affects the internal energy of the system, which effect increases as there are more disturbing B atoms dissolved. A point is reached when the energy contents of the system as a whole will be less if the atoms of B are separated and arranged on a space lattice of their own. The solid solubility of metal B in A will be as a first approximation inversely proportional to the amount of distortion of the normal space lattice of A caused by the substitution of an atom of B for an atom of A in that lattice. It is the power of the crystal of resisting slip which determines the hardness or power of resisting permanent deformation. "If this view is correct, then anything which tends to hinder the free occurrence of slip on the principal planes of a crystal will increase the hardness and will lower the ductility of the material."[2] Since a distortion of the lattice as mentioned above must serve as a hindrance to slip on the crystal planes, we should expect to find that the *hardening effect of one metal upon another in the form of a solid solution, should be inversely proportional to its solid solubility.* An example for low solubility is carbon in gamma iron (limit about 2 percent by weight). Accordingly the effect of carbon on the hardness of iron is great as compared with the effect of the highly soluble elements such as nickel or silicon.

(b) *The slip interference theory of Jeffries and Archer*[3]. In addition to a theory of hardness of solid solutions which is similar to that given by Rosenhain, the authors offer a theory to explain the hardness in general. They suggest that any structural condition which interferes with slip along the cleavage planes of the crystals of the metal increases its hardness.

(1) In a pure metal the most simple source of increased hardness is grain refinement, which introduces slip interference at the grain boundaries due to different orientations of the adjacent grains, and especially in fine grained metal between the grains.

(2) Cold working introduces slip interference by the fragmentation of the grains.

(3) Slip within grains is opposed by presence of hard grains uniformly distributed as fine particles.

(4) The effect of a given amount of hard constituent increases with the fineness of subdivision reaching a maximum at an average particle size denoted by the term critical dispersion (order of magnitude 10^{-7} cm). A higher degree of dispersion, particularly the atomic dispersion of solid solutions, is less conducive to hardness. Regarding the special case of

[2] Ewing and Rosenbain, Phil. Trans. A **249,** (1899)

[3] Zay Jefferies and R. S. Archer, Chem. Met. Eng. p. 1057, June 15, 1921 and p. 242, Feb. 8, 1922

steel, the authors conclude that martensite (structure of hard steel) consists of alpha iron which may hold carbon in solid solution. They claim that hardness is due chiefly to the grain refinement of the ferrite (alpha iron), but partly to the carbon in solution.

Object of the Present Investigation

It seems impossible to furnish by the microscope alone sufficient evidence to ultimately decide upon the various theories on the structure of martensite. The microscope fails to show the distribution of the carbon, the presence of which is essential to give iron the quality which allows it to be hardened. To bring the problem nearer to its final solution x-ray crystal analysis was resorted to by A. Westgren,[4] who in a pioneer investigation showed that martensite has the same body centered cubic structure as the alpha modification of pure iron. He found even the lattice parameters of the two to agree within the limit of observation. The only difference was that the x-ray spectral lines of the martensite were broader and less intense. Since very fine grain colloids give wider spectrum lines Westgren concluded that martensite may consist of alpha iron grains which are of about the same size as the grains of gold colloids. Undoubtedly such a fine grain size would also tend to explain the hardness of martensite.

Since, however, no martensite can be produced in absence of carbon, it seems that carbon must play an essential role in rendering steel hard. It was, therefore, of importance to trace the carbon in the iron lattice of hard steel. Due to both its low atomic number and the relatively small quantity present, it seemed to be hopeless to do this directly. It was, however, thought possible that from small changes found in the x-ray spectrum indirect conclusions as to the distribution of carbon might be drawn.

Evidently a method which would show the distribution of intensity of the x-ray spectrum, such as the ionization chamber, would be more promising for this purpose than the photographic method. It was, therefore, decided to study the structure of martensite under different conditions by means of an x-ray spectrometer employing the ionization chamber as an indicator.

The brick shaped specimen was placed on a small spectrometer, with its polished and slightly etched plane through the axis of rotation of the tables. The smallest interval of reading on the vernier is 5′ of arc on the chamber slit, and 10′ on the specimen table. The arrangement of

[4] A. Westgren, J. Iron and Steel Inst., p. 303, 1921; and London Engineering, May 19, 1922

slits and target is shown in Fig. 1. The slits, axis of spectrometer and target were lined up with a silk thread. All the slits are cut into sheet lead of 3 mm thickness and are non-adjustable. The ionization slit, 1 mm wide and 5 mm high, is mounted on an arm which is rigidly connected with the table.

The ionization chamber is mounted on a support concentric with the spectrometer and resting on a reinforced wooden board to relieve the spectrometer from carrying its weight. This chamber consists of an earthed brass tank 40 cm long by 10 cm in diameter filled with methyl bromide vapor. All joints and pet cocks were sealed with shellac. The leak was checked by ionization measurements and was found to be negligible even after a months time. The methyl bromide was specially prepared and was passed over phosphorous pentoxide and soda lime before

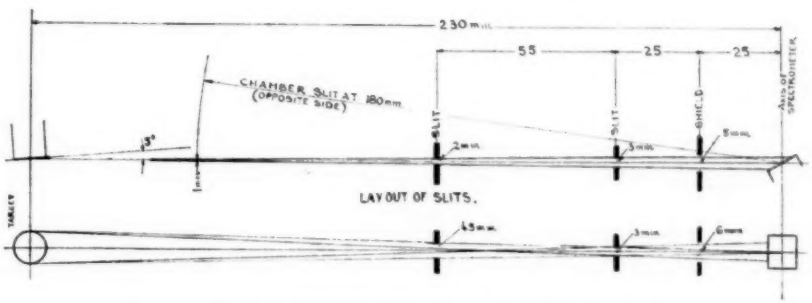

Fig. 1. Diagram showing path of x-rays through apparatus

being admitted in the tank. Inside this tank was a 3 mm diameter brass rod electrode insulated by amber, and surrounded by another electrode which is a brass cylinder of oval section. The saturation voltage on the plate was found to be about 40 volts at maximum vapor density and strongest ionization occurring during the experiments, hence 72 volts was considered safe to maintain saturation under all conditions. There are two openings 1 cm in diameter sealed by thin mica windows, one at each end of the chamber, enabling a convenient lining up of the chamber with the slit. Care was taken to prevent the beam from striking the electrodes. The wire connecting the chamber with the electrometer is shielded by a sealed brass tube of 25 mm diameter.

The electrometer is of the Dolezalek type, with a sensitivity of 3500 mm per volt. On the advice of Dr. K. T. Compton a platinized mica needle of only 22 mm diameter was used, which reduced the period of swing of the needle at a given sensitivity. This was suspended by a fine Wollaston wire.

The angular speed of the needle at different intensities of ionization was observed at the different parts of the scale and at each intensity it was found to be constant within one or two per cent at any reading beyond 40 mm. Readings were thus started at 50 mm from zero.

Considerable trouble was experienced with the earthing key and all the different types tried failed after a certain number of breaks. A cleaning and sharpening of the contacts removed the trouble only for a limited time.

The x-ray tube was of the well known water-cooled molybdenum target Coolidge type. It was enclosed in a lead box with walls 1/4 in. thick. The x-ray generating outfit and the measuring systems were installed in adjoining rooms, the systems being separated by a lead door 1/8 in. thick. Measurements comparing ionization with the x-ray tube turned off and then on, but with specimen removed from the spectrometer showed that shielding from stray rays was perfect.

The x-ray tube was directly connected with the secondary of the transformer, the tube rectifying its current. The tube voltage was kept constant by a small variable resistance and a Weston A.C. voltmeter, both inserted in the primary of the transformer. The absolute peak value of the secondary voltage was determined by the vanishing point of the x-ray spectrum at the short wave end and was found to be 42,000 volts at 25.3 m-amp. tube load.

The Specimens

The size of all hardened specimens was $1.2 \times 1 \times 0.4$ cm.

Specimen 1 (*carbon free iron*). Iron melting bar. This sample was almost free from carbon, but contained a considerable amount of silicious slag as seen under the microscope. It was not analyzed. To refine the grain, it was heated to 900°C, held for six minutes, quenched in oil. Treated on March 1st, 1923.

Specimen 2. Analysis: C = .80%, Cr = .14%, Mn = .35%, Si = .19%. Other ingredients were not determined. *Treatment 2f* (*fine martensite*). Quenched three times from 850°C in oil to dissolve the cementite (Fe_3C). It was then heated to 765°C to give fine martensitic structure; quenched in ice water to avoid troostite. The microscope at 1000 diameters magnification shows absence of cementite. The martensite needles are very fine. No austenite or troostite could be seen. *Treatment 2c* (*coarse martensite*). Quenched three times from 850°C in oil; final quench in ice water from 1030°C to give coarse structure. Microscope at 1000 diameters shows absence of cementite, coarse martensite needles and no austenite

or troostite. *Specimen 2a* (*annealed cementite*) was cut from the original annealed bar containing globular cementite.

Specimen 4. Analysis: C = 1.31%, Cr = .12%, Mn = .24%, Si = .17%. Other ingredients were not determined. *Treatment 4 f* (*fine martensite*). Quenched three times from 850°C and twice from 1035°C in oil to dissolve cementite; final quench from 850°C in ice water. No austenite or troostite could be seen under the microscope, and but little undissolved cementite; the martensite was fairly fine. *Treatment 4c* (*coarse martensite*). Quenched three times from 850°C and twice from 1035°C in oil; final quench from 1035° in ice water. No austenite or troostite could be seen under the microscope, and only a few fine grains of cementite; the martensite was coarse. In order to avoid decarburization the specimens were soaked for only about two minutes after their surface assumed final temperature. Specimens 2f, 2c, 4f, 4c, were heat treated at end of Aug. 1923. *Specimen 4a* (*annealed cementite*) was cut from the original annealed bar containing globular cementite.

TABLE I

Hardness and density of the specimens.

Specimen	Brinell hardness	Density
2a(0.8% C) annealed	140 kg/mm	7.823 g/cm³
2f fine martensite	660	7.757
2c coarse martensite	640	7.737
2 fine martensite drawn to 260° C for 30 minutes	590	7.781
4a (1.31% C) annealed	170	7.793
4f fine martensite	—	7.729
4c coarse martensite	640	7.729

DATA OBTAINED WITH THE X-RAY SPECTROMETER

Readings were taken in at least 30′ intervals over a range of 36° to 50°. Fig. 2 shows the ionization current as a function of the slit angle for specimens 1 and 4f. The reading corresponding to the primary beam was at 45°.

The lines with the corresponding atomic spacings of the carbon free iron are compiled in Table II. The location of lines check with those found by other investigators within close limits and the space centered lattice of alpha iron is once more verified.

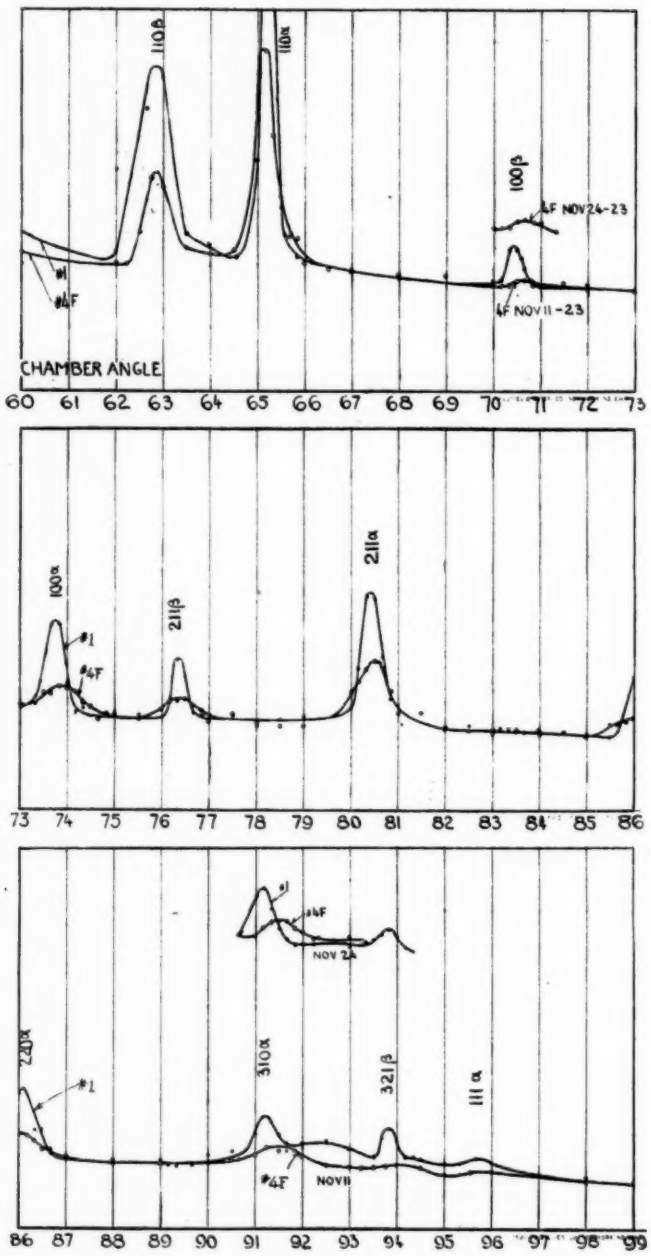

Fig. 2. X-ray spectra for specimens #1 and #4F; ionization as a function of angular position of the slit. Peak marked 220α should be 110α(2).

TABLE II

X-ray spectrum of carbon free iron

Reading	θ	Line*	Spacing	a	Deviation
62° 50′	8° 55′	110β	2.0291A	2.870A	+0.21%
65° 10′	10° 5′	110α	2.0305	2.872	+0.28
70° 25′	12° 42′	100β	1.4296	2.859	−0.14
73° 45′	14° 22′	100α	1.4319	2.864	0
76° 20′	15° 40′	211β	1.1646	2.854	−0.35
80° 30′	17° 45′	211α	1.1661	2.858	−0.21
86° 5′	20° 32′	110α(2)	1.0131	2.865	+0.03
91° 12′	23° 6′	310α	0.9061	2.866	+0.07
93° 50′	24° 25′	321β	0.7608	2.847	−0.59
95° 40′	25° 20′	111α	0.8308	2.880	+0.56
100° 20′	27° 40′	321α	0.7656	2.865	+0.03
108° 30′	31° 45′	{411α {110α(3)	0.6755	2.865	+0.03
			Mean	2.864†	

* The line whose peak is measured is designated by the plane and the wave-length $\alpha = K\alpha = .711A$; $\beta = K\beta = .629A$

† The density of extremely pure vacuum fused iron was recently (April 1924) determined by the Bureau of Standards to be 7.8685 at 20°C. This gives for α the value 2 8691A, 0.17 per cent greater than the above mean. The metal had been rolled and contained: C 0.001%, S .011%, Si .004%, Mn none, P not determined. These results are published with the kind permission of the director Dr. George K. Burgess.

The main object of taking the spectrum of α iron was, however, for comparison with that of martensite. It was found that martensite shows the same lines in number and location as soft iron. The most outstanding difference is the decrease of intensity of the martensite peaks to about one third of the intensity of the iron peaks, and their slight broadening which is more pronounced at the right end of the spectrum. Also, particularly at the right end of the spectrum, a shift of the martensite lines to the right was noticed, indicating a contraction of the martensite lattice compared with that of pure iron. The formula for the relative value of contraction may be obtained by differentiating Bragg's formula $2a \sin \theta = \lambda$, giving for small increments,

$$\Delta a/a = -\Delta\theta\sqrt{(2a/\lambda)^2 - 1}$$

To determine the shift more accurately, peak 310α was gone over in detail taking many readings at close intervals, and repeating each reading two to four times (see Fig. 2). The martensites with different carbon analyses seem to show approximately the same amount of shift. This indicates that only a limited amount of carbon can be dissolved (which is in accordance with Rosenhain's view quoted above). The average value of the shift of six determinations was found to be $2\Delta\theta = 14'$. The relative linear contraction is therefore $\Delta a/a = -.00469$; and the relative volume contraction is $\Delta v/v = 3\Delta a/a = 0.0141$. 211α was treated in a similar way and gave a shift of only about 5′ and in the opposite direction

to that for 310α, which the writer is unable to interpret at the present time.

Outside of the α iron lines one feeble austenite line at 67°50′, $\theta = 11°25'$, could be seen on specimen 4c (1.31% carbon, coarse martensite).

The spectrum of the 1.31% carbon steel was also determined in the annealed condition. It is interesting to note that the intensity of the lines is somewhat lower than in case of pure iron. This is due to the large amount (about 20%) cementite present, which displaces the iron, and consequently reduces the quantity of scattering substance. A few cementite lines while weak were readily identified. Spectra of coarse and fine martensite of medium and of high carbon steel showed no marked differences.

Structure of Martensite Derived from the Foregoing Data

The x-ray spectrum of the martensite determines its space lattice to be a body centered cube as in the case of α iron. The shift which was determined on one line only, shows that the martensite lattice is spaced approximately 0.4 per cent closer than the α iron lattice. There also is a decrease of intensity of the martensite lines to about 0.4 that of pure iron lines, accompanied by slight broadening.

This evidence together with the densities should be considered in connection with the following various possible assumptions:

(a) All carbon is present as extremely fine grains of cementite (Fe_3C), too small to be visible under the microscope. In this case no distortion of the α lattice would be expected.

(b) The lattice is substitutional, the carbon replacing the iron atoms. This would cause a drop of intensity of the spectrum lines due to the distortion of space lattice by the comparatively small carbon atoms.

(c) The carbon is within the lattice. This also may cause distortion and might also contract the lattice.

In order to be able to choose between the assumptions enumerated the densities to be expected in the case of each of these assumptions will be determined.

(A) *The density of a mixture of cementitie and α iron.* If the mass of carbon combined as Fe_3C is x, and the ratio of the atomic weights is $p = Fe/C$ it can readily be shown that, taking the density of cementite[5] to be 7.59 the density of steel containing all the carbon as cementite is

$$\frac{1}{\rho} = \frac{x}{7.59}(3p+1) + \frac{1}{7.86}\Big(1 - x(3p+1)\Big).$$

[5] Levin and Dornhecker, Ferrum **10**, 321 (1914)

(B) *The density of a substitutional solid solution of the body centered lattice type.* In the case of a body centered lattice, there are two atoms per elementary cube. Since some of the atoms of one kind are substituted by the other kind we may speak of an average number n_C, n_{Fe} of atoms of each kind per cube; then

$$n_C + n_{Fe} = 2 \ .$$

If the proportional weight of carbon present is y, then the weight ratio of carbon to iron is

$$y/(1-y) = n_C C / n_{Fe} Fe \ .$$

Hence the average number of iron atoms per elementary cube is

$$n_{Fe} = \frac{2(1-y)}{x(p-1)+1} \ .$$

Let ρ denote the density of the solid solution of carbon in iron, mC the absolute mass of a carbon atom and a the lattice parameter, then

$$\rho = \frac{n_{Fe} Fe + (2 - n_{Fe}) C}{a^3} = \frac{2pmC}{a^3} \cdot \frac{1}{x(p-1)}$$

(C) *The density of the martensitic steel* under the assumption that all carbon is inserted and forms no part of the structure, is

$$\rho = 7.86\ (1+y)$$

provided the lattice is unaltered. Since the x-ray spectrum shows that the lattice contracts, the density would be about 1.014 times greater.

The densities calculated under the various assumptions are as follows for 0.8 and 1.31 per cent C:

	.80% C	1.31% C
(A) Martensite is a mixture of submicroscopic Fe_3C and Fe	7.825	7.803
(B) Solid solution, all carbon substituting the iron atoms, lattice assumed to be uncontracted.	7.637	7.490
(B′) Solid solution, all carbon substituting the iron atoms, lattice contracted 0.4%	7.74	7.60
(C) Solid solution, carbon between the lattice, lattice assumed the same as that of pure iron.	7.923	7.963

In Fig. 3 the densities calculated under these various assumptions are shown graphically as a function of the relative weight content of carbon. The actual densities of the different specimens as determined by weighing are also entered in the diagram. These data show that a considerable part of the carbon present in martensite replaces iron atoms in the space lattice. Due to the fact that the carbon atoms are in the proportion 5:9 smaller in diameter than the iron atoms which they replace, there must be considerable lattice distortion. But a distorted lattice must scatter x-rays very like a lattice subjected to temperature agitation. I. Backhurst[6] found that the peak intensity due to a single crystal of aluminum heated

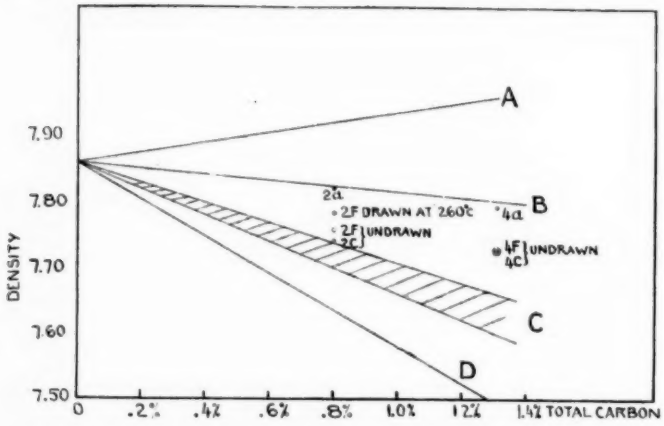

Fig. 3. Density as a function of per cent carbon for various assumptions.
A: All carbon in solution and assumed within the lattice; (C)
B: All carbon present as cementite; (A)
C: Carbon replacing iron atoms in lattice, causing contraction; } (B)
D: Carbon replacing iron atoms in lattice, without contraction. }

to several hundred degrees Celsius is considerably less than the cold value, the relation of the spectra of cold and hot aluminum thus being similar to the relation between the spectra of pure iron and martensite. It is, therefore, probable that martensite is a distorted lattice of iron. This would also explain its hardness.

The assumption of colloidal crystal grains would also account for the drop of intensity of the martensite lines. In this case the lines corresponding to the planes of lowest atomic population (fewest number of lines in the grating) should drop to the greatest extent in intensity and vice versa. This is contrary to what was observed, the drop of line (100) being the greatest.

[6] I. Backhurst, Proc. Roy. Soc. A **102**, 340 (1922)

Summarizing the observed facts it seems probable that *martensite is a mixture, consisting of two constituents which may be present in varying proportions, one a solid solution of carbon in α iron, the carbon atoms replacing the iron atoms, the other finely dispersed cementite.*

The solution is unstable and some of the carbon gradually precipitates from the lattice even at room temperature; hence it is difficult to keep more than part of the carbon in solution.

The gradual unloading of carbon from the lattice at room temperature is shown in Fig. 4 where the 310α line of two specimens is compared, both made of the same steel, treated alike, one about five months, the second only three or four days old. The intensity of the lines of the former is

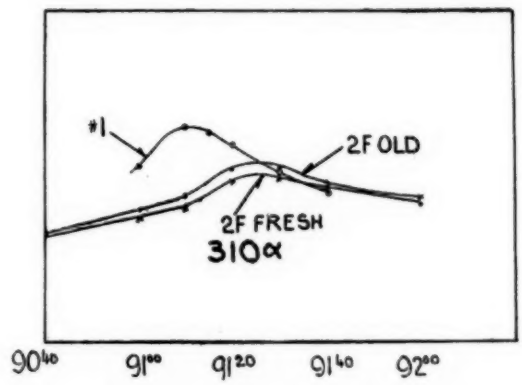

Fig. 4. Influence of seasoning on the line 310α.

slightly greater. To prevent possible error due to slight change of sensitivity of the electrometer, readings were taken in succession on the three specimens shown in Fig. 4, before changing the setting of the spectrometer table.

Fig. 3 shows that there is a limit of the amount of carbon held in this unstable solution. Specimens 4f and 4c have considerably higher density than the total carbon in solution would require. Hence the balance is likely to be present as dispersed submicroscopic cementite.

Fig. 5 shows a few drawing experiments. The degree of lattice distortion indicated by the intensity of lines is small after drawing to a temperature of 260°C applied for thirty minutes. The density of this specimen increased and approached the all cementite line (Fig. 3). The fact that the cementite cannot yet be seen under the microscope may be due to the submicroscopic size of the Fe_3C grains, but their presence was proved by the well known darkening of the polished surface by etching in weak acids.

It is interesting to note that at the drawing temperature of 260°C the lattice distortion has to a great extent disappeared; on the other hand the density in Fig. 3 shows that only a part of the carbon has been precipitated. The hardness only dropped from 660 to 590 kg/mm², which is far greater than the hardness of the annealed state, 140 kg/mm². It may be that in the undrawn martensite hardness is chiefly due to lattice distortion as required by Rosenhain's theory of solid solutions, but almost the same hardness may be maintained by the carbon which is precipi-

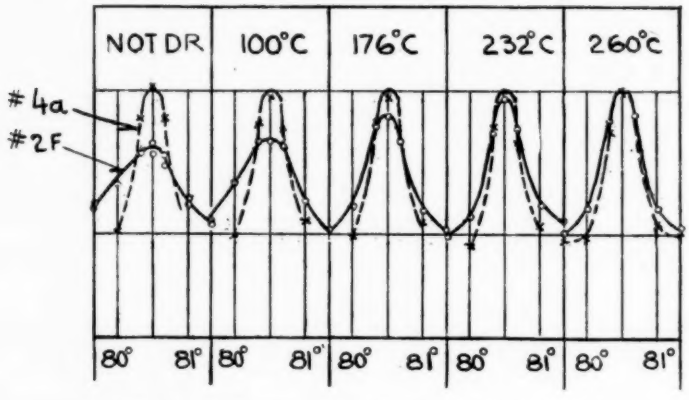

Fig. 5. Influence of drawing on the line 211α.

tated at drawing to a moderate temperature, say at 260°C and is likely to be present as cementite in fine dispersion, when the distortion of the lattice is negligible (Jeffries and Archers slip interference theory).

In concluding, I wish to express my indebtedness to the S K F Industries Inc. for giving me the privilege of devoting part of my time to this research, to Prof. C. B. Bazzoni, Dr. Lucian of the University of Pennsylvania, to Dr. H. Styri and Mr. Frank Wright of the S K F Industries, to Prof. C. W. Hewlett and to Prof. K. T. Compton.

RESEARCH SECTION OF THE RANDAL MORGAN LABORATORY,
UNIVERSITY OF PENNSYLVANIA,
May 18, 1924.

THE BARKHAUSEN EFFECT

By E. P. T. Tyndall

Abstract

Barkhausen effect for silicon steel.—In a typical experiment a strip of silicon (4.2%) steel, after having been carried through several hysteresis cycles and brought to a steep part of the *B-H* curve, was subjected to a magnetic field which was increased in a continuous manner by 0.13 gauss in 2 sec. To obtain a record of the discontinuities in magnetization, the specimen was surrounded by a small search coil connected through an amplifier to a moving coil oscillograph. The *oscillograph records* show many almost instantaneous deflections with a random distribution both as to time of occurrence and as to magnitude, each followed by an approximately exponential return to zero. The duration of an impulse depends probably on eddy currents in the specimen. A rough time constant of 3×10^{-5} sec. is computed for a 4% silicon steel wire, 1 mm in diameter. The apparatus was calibrated by means of artificial impulses of the same type. Assuming each impulse is due to a sudden saturation of a small portion of the material, the change in magnetic moment of this portion is found to vary from .001 to .008 e.m.u.; the average change .003 is sufficient to saturate a volume of 1.7×10^{-6} cc. This volume, while comparable with that of a crystal grain in the particular case described, was found not to depend upon grain size in any marked degree in other experiments. The results favor the suggestion of Barkhausen that magnetic materials magnetize discontinuously, but leave open the question as to what determines the size and shape of the portions which suddenly change.

Introduction

THE discontinuities in the magnetization of iron and other magnetic materials have been investigated by Barkhausen,[1] Van der Pol,[2] Gerlach and Lertes,[3] Weiss and Ribaud,[4] Zschiesche,[5] and S. R. Williams.[6] Barkhausen and Van der Pol consider the effect as due to the sudden reorientation of groups or chains of molecular magnets, possibly all the magnets in a single crystal turning simultaneously. The latter has published pictures of the discontinuities as recorded by a fluxmeter and by a string galvanometer. Gerlach and Lertes, and Zschiesche placed the specimen in a solenoid which was then rotated at an arbitrary rate between the poles of an electromagnet. The amplified and rectified

[1] H. Barkhausen, Phys. Zeits. **20,** 401 (1919)
[2] B. Van der Pol, Proc. Acad. Amst. **23,** 637, (1921); **23,** 980 (1922)
[3] W. Gerlach and P. Lertes, Zeits. f. Phys. **4,** 383 (1921)
[4] P. Weiss and G. Ribaud, Jour. de Phys. (6) **3,** 74 (1922)
[5] K. Zschiesche, Zeits. f. Phys. **11,** 201 (1922)
[6] S. R. Williams, Phys. Rev. (2) **22,** 526 (1923)

current from the solenoid as measured by a galvanometer is defined by them as the "effect." Various experiments led them to the conclusion that the effect is dependent on magneto-striction and not primarily on discontinuities in magnetization. Williams has recently obtained oscillograms of the effect using a telephone receiver as an oscillograph. All observers seem agreed that the effect takes place only where the *B-H* curve for the material is steep. Preliminary work in this laboratory showed that the effect as measured by the noise in a telephone receiver is roughly proportional to the slope of the *B-H* curve. Iron dust, loose or compressed, gave a negligible effect. The work reported here was undertaken to test the hypothesis that the phenomenon is due to the magnetization of single crystals and to obtain a more detailed picture of it than has yet been presented. Most of the work has been done on one specimen, though enough observations have been made on others to give confidence in the generality of the results.

The discontinuities in magnetization are best measured by the induced voltages in electrical circuits, magnetometer methods being unsuited to the proper analysis of the effects. In an attempt to record the discontinuities separately and with as little distortion as possible various methods have been tried. The first instrument used was the oscillograph tube developed in this laboratory by J. B. Johnson.[7] This gave visual evidence of the discontinuities but it was at once apparent that photographs could not be taken quickly enough to follow the changes in a single event. Later a moving coil oscillograph of the ordinary type[8] was used and found more satisfactory.

Experimental Method

The circuit used in obtaining oscillograms of the effect is shown in Fig. 1. Changes in magnetization are produced by passing current through either of two single-layer solenoids S_1 and S_2, each 80 cm long and 2.5 cm in diameter, wound in parallel on a glass tube. The search coil S_3 is a multiple-layer solenoid of 10,000 turns, 2.5 cm long and 2.2 cm in outside diameter. The specimen is a straight wire or flat strip longer than the search coil and supported at the center of the magnetizing coils S_1 and S_2. The voltage generated in S_3 by sudden changes in magnetization of the specimen is amplified by the amplifier A, the output of which goes to the oscillograph (Vibrator No. 2). The circuit connected to S_1 is used to carry the specimen through several magnetizing cycles

[7] J. B. Johnson, Jour. Opt. Soc. Am. and R.S.I. **6**, 701 (1922)
[8] General Electric Company Oscillograph, Type EM Form C.

before taking an exposure. The specimen can thus be brought to any desired point in the cycle. The experimental change in magnetization is then effected by closing the filament circuit of the vacuum tube V. As the temperature of the filament rises the current through S_2 grows continuously and almost linearly with time, thus giving a smooth variation of magnetizing field. This current is recorded on the oscillograph (Vibrator No. 1). A time scale is given by a 60-cycle alternating current (Vibrator No. 3). No transformers are used in the amplifier circuit, the coupling being by means of resistance (100,000 ohms) and capacity (0.1 mf). A "V" tube[9] is used in each of the first three stages; in the fourth are two "L" tubes,[10] connected in parallel. The space current (20×10^{-3} amp.) from the last stage passes directly through the oscillograph vibrator. An

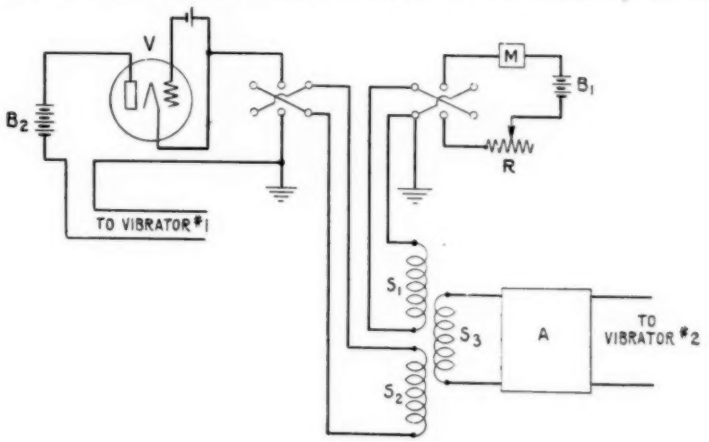

Fig. 1. Diagram of apparatus for recording Barkhausen effect.

amplifier of this type is comparatively free from distortion. The effect on the records due to amplifier and oscillograph characteristics is discussed below.

Before taking an oscillogram the specimen is taken about ten times around a hysteresis loop wide enough to ensure nearly complete saturation at either end, and so brought to some desired point, usually that corresponding to zero current in solenoid S_1. The magnetizing field thus produced is not quite zero since the vertical component of the earth's field is not neutralized. The light is then screened from vibrators No. 2 and No. 3 and the zero position of vibrator No. 1 marked on the film during one rotation of the drum. After these preliminaries the oscillogram is taken at once to minimize the disturbing effects of stray magnetizing

[9] Western Electric Company, 102-DW
[10] Western Electric Company, 216-A

fields. The exposure is made just after closing the filament circuit of the auxilliary vacuum tube V. To simplify the comparison of records two standard rates of the oscillograph drum were chosen. In the slower case the closing of the key in the filament circuit of V also opens the shutter of the oscillograph by means of a relay. The picture thus obtained shows the process from the beginning nearly to the end, lasting about two seconds. In the faster records the vacuum tube filament is lighted and after a suitable short interval the shutter is opened. A record at this rate covers half a second or less and a more detailed picture is given of a small part of the process.

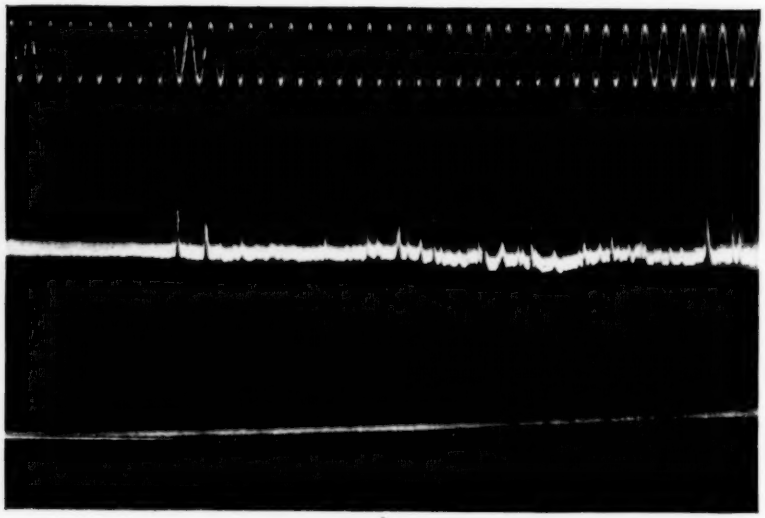

Fig. 2. A record taken at the slower rate.

Data and Discussion

A portion of a typical film, taken at the slower rate, is shown in Fig. 2. This record was obtained with a strip of silicon steel (4.2%) 0.043 cm × 0.178 cm × 17.8 cm which had been annealed in vacuum for 11 hours at 900°C and had then had a further heat treatment, also in vacuum, of 30 min. at 1200°C. The average linear dimension of the grains as measured on etched areas is about 1.5×10^{-2} cm. The record of Fig. 2 shows well the discontinuities in the process of magnetization. The amplified curve is approximately the derivative of the B-H curve for the specimen. It shows that a portion of the B-H curve consists of short straight lines parallel to the B axis and H axis alternately.

It might be expected that consecutive experiments under identical conditions would give superposable oscillograms, the magnitude and order of the impulses being completely determined by the initial state of magnetization. In no case, however, was it possible to obtain two identical records, even though many successive films were compared, in which the specimen had been left undisturbed in position and had been brought to the same state through identical preliminary magnetizing cycles.

A study of the occurrence and magnitude of the discontinuities was made. Five successive records like that of Fig. 2 were analyzed. The change in the magnetizing field was 0.130 gauss and the change in the induction corresponding thereto was about 2000. This increment was divided into intervals 0.005 gauss long and the number of impulses in each interval recorded. A count was then made of the number of intervals containing 0, 1, 2, 3 etc. impulses. These data are graphically presented in Fig. 3. The ordinate of any point is the number of field intervals

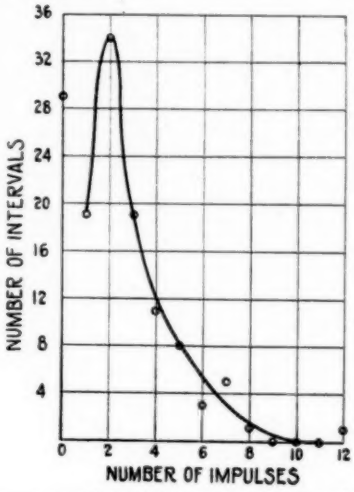

Fig. 3. Distribution among equal field intervals of discontinuities in magnetization.

containing the number of impulses shown by the abscissa. The graph shows that this distribution is a random one, except for the excess of intervals containing no impulse, most of which occur at the beginning of the record. It is thought that this results from the unavoidable fluctuations in magnetizing field between the end of the preliminary treatments and the beginning of the exposure. In Fig. 4 is plotted the distribution of the impulses as to magnitude. The ordinate of any point is the number of impulses having a magnitude (simply the deflection in cm measured on the film) of the size given by the corresponding abscissa. This distribution also appears to be a random one.

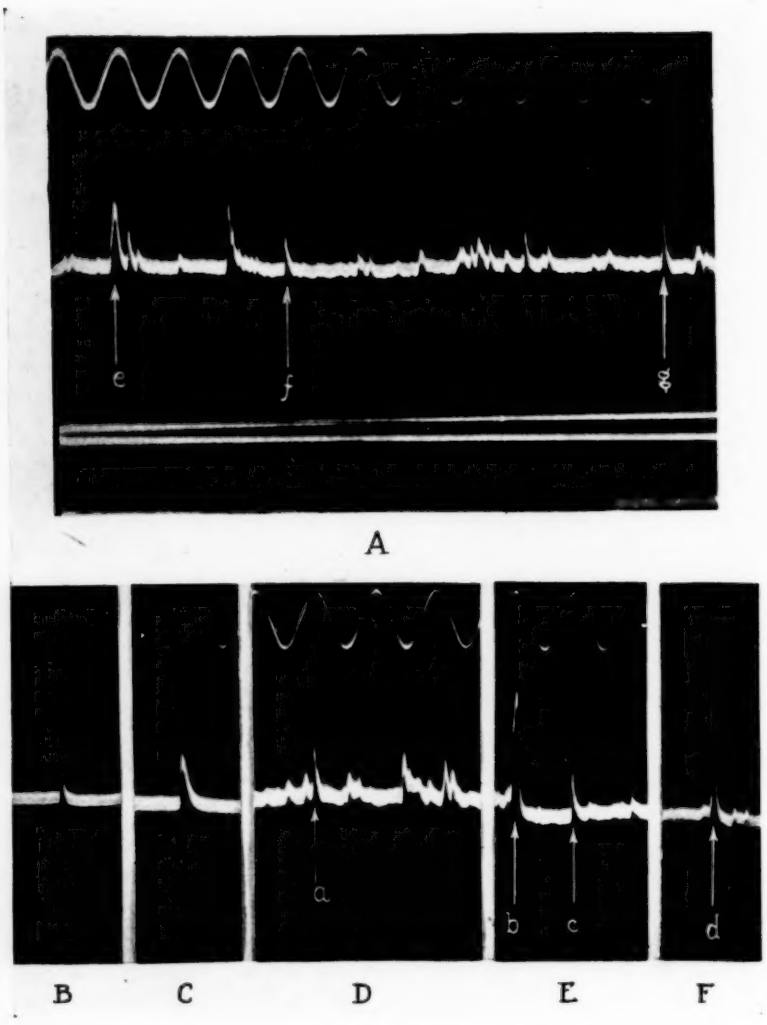

Fig. 5. Records taken at the faster rate.
A. Change in H 0.022 gauss. Measured impulses e,f,g.
B. Artificial impulse, $\tau = 0.1 \times 10^{-3}$ sec.
C. Artificial impulse, $\tau = 1.45 \times 10^{-3}$ sec.
D, E, F. Measured impulses a,b,c,d.

In a record of the second kind described above each impulse consists of an almost instantaneous deflection followed by an approximately exponential return to zero. Part of one such record is shown in Fig. 5A. Oscillograms taken with fresh damping fluid show higher and sharper peaks, corresponding to the lower value of damping factor thus obtained.

To calibrate the amplifier artificial impulses were similarly recorded by the arrangement shown in Fig. 6. The search coil S_3, and the amplifier circuit are as before, there being, however, no specimen in S_3. In the circuit with S_2 are now an inductometer (L_3, r_3), resistance r_4, and milli-

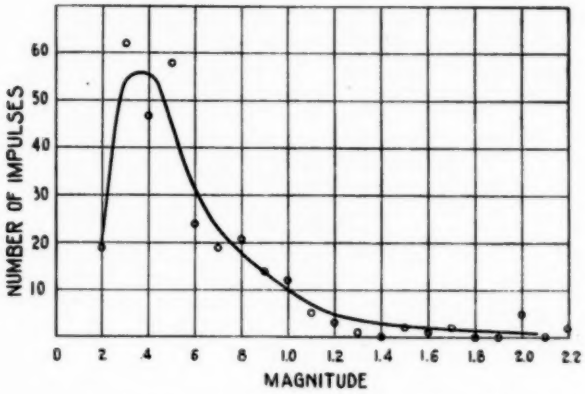

Fig. 4. Distribution in magnitude of discontinuities in magnetization.

ammeter M. By adjusting the ratio of r_4 to R and the value of L_3 it is possible to establish a current of a chosen magnitude in the circuit in which S_2 is included and to adjust the time constant of this circuit as desired. The equation for this current is given by

$$i = I(1-\epsilon^{-t/\tau}) \tag{1}$$

in which

$$\tau = (L_2+L_3)/(r_2+r_3+r_4)\ . \tag{2}$$

In these expressions I is the steady state current with the key K closed and τ is the time constant. The voltage induced in the search coil S_3, is

$$e = M di/dt = (MI/\tau)\epsilon^{-t/\tau} \tag{3}$$

where M is the mutual inductance of S_2 and S_3. It is an amplified picture of this voltage which is recorded on the film as a function of the time. Its initial value and the area under the curve are given by

$$\epsilon_m = MI/\tau \quad \text{and} \quad \int_0^\infty e dt = MI\ .$$

Oscillograms showing two such impulses are shown in Fig. 5B, 5C. Table I gives results of measurements on such artificial impulses, the deflection h, and the area under the curve A, being the quantities measured from the oscillograms.

The values of the ratio e_m/h are interesting in that they show how the amplifier and oscillograph distort a very sudden voltage impulse, the maximum value being cut down while the rate of return is diminished. If there were no distortion the ratio e_m/h would be constant. In spite of the distortion, however, the system is a competent integrator, for, as the

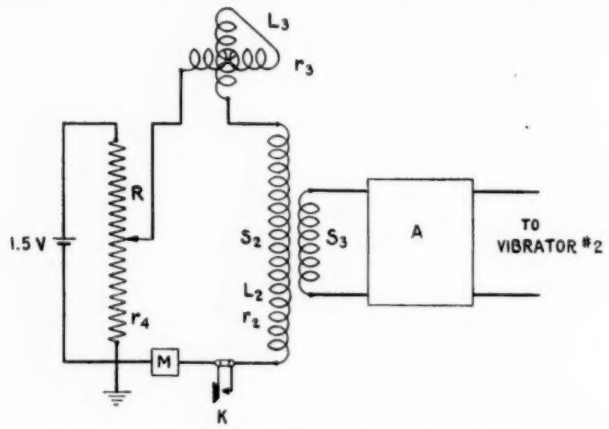

Fig. 6. Diagram of apparatus for producing and recording artificial impulses.

last column shows, $\int edt/A$ is approximately constant for values of τ down to 10^{-4} seconds. It is moreover considered constant for values of τ less than 10^{-4} seconds, as it was later checked experimentally to $\tau = 10^{-5}$

TABLE I

τ	e_m	$\int edt$	h	A	e_m/h	$\int edt/A$
($\times 10^{-3}$)	($\times 10^{-3}$)	($\times 10^{-6}$)		($\times 10^{-3}$)	($\times 10^{-3}$)	($\times 10^{-3}$)
0.1 sec.	2.93Volts	0.293	0.75 cm	0.64	3.9	0.46
0.1	6.30	0.630	0.78	0.97	8.1	0.65
0.2	3.21	0.648	0.90	0.65	3.6	1.00
0.2	6.25	1.25	1.33	1.26	4.7	0.99
0.2	1.53	0.31	0.55	0.635	2.8	0.49
0.2	3.97	0.80	1.02	1.30	3.9	0.62
0.28	8.16	2.28	2.13	3.55	3.83	0.64
0.3	3.11	0.923	1.17	1.64	2.66	0.56
0.41	2.85	1.17	1.58	1.94	1.8	0.60
0.60	1.80	1.08	1.17	1.64	1.5	0.66
0.62	3.68	2.28	1.86	3.12	2.0	0.73
0.63	1.79	1.13	1.60	2.42	1.1	0.47
1.45	0.764	1.11	1.01	1.78	0.76	0.62
					Average	0.652

seconds, there being no considerable change in $\int edt/A$ between $\tau = 10^{-3}$ and $\tau = 10^{-5}$ seconds. These data are not included in the table as they were taken after replacement of the damping fluid in the oscillograph, which changed the actual value of $\int edt/A$ for any value of τ.

From oscillograms of the effect taken at the faster rate some speculations concerning the magnitude of the change in magnetic moment producing a sudden voltage impulse may be made. Impulses of various sizes, all from the same specimen previously described and with zero current in S_1, Fig. 1, were studied. Some of these are shown in Fig. 5 already referred to. The time constant of any of these impulses is by inspection not more than 10^{-3} seconds and may be very much less. Since it has been shown that the area under the curve of the amplified voltage is practically independent of the time constant of the voltage impulse, an estimate may be made of the total change in magnetization corresponding to any impulse without knowing the time required for the change. It will merely be assumed that a change in magnetization of a small portion of the material takes place in a short space of time. The portion of material will be considered as a dipole whose magnetic moment changes suddenly from one value to another. Let M be the variable magnetic moment of the dipole and consider it situated at the center of the short solenoid S_3, Fig. 1. It is then very easy to show that the area under the curve for the voltage induced in such a solenoid when M changes from 0 to M is

$$\int e dt = \frac{2\eta M n}{10^8 \sqrt{a^2+R^2}} \text{ (volts–sec.)}$$

where n is the number of turns on the solenoid S_3, a is half its length in cm, and R is the mean radius of the multiple layer winding. Some objection may be made to the assumptions that have been made but until more is known of the mechanism of magnetization they do not seem unreasonable. The dipole has been considered as surrounded by non-magnetic material. This assumption, however, does not seem so arbitrary when it is remembered that it is evident from the records that the rest of the material does not greatly change its magnetization when one small portion does. That a sudden change in magnetization might set up eddy currents in the surrounding material which would modify the observable effect, is quite possible, and will be discussed later.

For a given Barkhausen impulse the $\int e dt$ may be determined from the area by the calibration factor derived in Table I. The experimental data on a series of impulses, shown in Table II, have been interpreted on the basis of the analysis given above. The letters a to g designate the impulses shown on Figs. 5. For simplicity the area under each deflection has been approximated to that of a triangle whose base b is the time during which the vibrator is deflected from the zero position.

TABLE II

Impulse	b	h	Area	$\int e dt$	M
	($\times 10^{-3}$)				($\times 10^{-3}$)
a	2.16 sec.	1.00 cm	1.08	0.71	1.5 e.m.u.
b	2.43	2.15	2.61	1.70	3.6
c	1.69	0.95	0.80	0.52	1.1
d	2.60	1.62	2.10	1.37	2.9
e	3.04	1.28	1.95	1.27	2.7
f	1.59	0.65	0.52	0.34	0.72
g	1.85	1.10	1.02	0.67	1.4
h	2.58	1.20	1.55	1.01	2.1
i	1.98	0.72	0.71	0.46	0.97
j	1.52	0.70	0.53	0.35	7.3
k	1.06	1.05	0.56	0.37	7.7
				Average	2.9

DISCUSSION OF RESULTS

Assuming that a small volume v of the material is magnetized to saturation to produce each of the changes in magnetic moment given in Table II, the average value of v is given by

$$v = M/\sigma .$$

Putting σ, the intensity of magnetization at saturation, equal to 1700 gives

$$v = 1.7 \times 10^{-6} \text{cm}^3 ,$$

and

$$v^{1/3} = 1.2 \times 10^{-2} \text{cm} .$$

Although this is not far from the average linear dimension of the grains in the specimen, it is not to be concluded that the grains become magnetized as units. Experiments to be quoted below will make this clear. However the change in magnetization arises, the order of magnitude of the change has been estimated on the basis of somewhat speculative assumptions.

It will be interesting to compare the result with that obtained by another experimenter. Van der Pol assumes groups of elementary magnets in long filaments and computes that the change in flux corresponding to one discontinuity is 0.33 lines. On the same assumption the average magnetic moment shown in Table II would correspond to a change in flux of 0.14×10^{-2} lines. Van der Pol used a different material and a different method and it is not surprising that the results should differ. An experiment was performed to determine whether the filaments which he describes existed in the specimen here studied. Two identical search

coils, one of which was that previously used, were placed end to end over the specimen and substituted in the measuring circuit for the single coil S_3, Fig. 1. Connections were such that a change in magnetization of given sign would cause oppositely directed voltage impulses in the two coils. The record showed deflections on both sides of the zero line, similar in type and approximately equal in number and magnitude to those obtained with a single search coil, clearly indicating that few changes in magnetic moment affected both coils. This was taken to mean that this sample differed from that studied by Van der Pol in having few long filaments or at least in having few that became magnetized as units. The result of this experiment justifies the assumption that the magnetic element may be treated as a dipole.

To test the effect of grain size on the phenomenon a strip of the same silicon (4.2%) steel sheet was vacuum annealed at about 1300°C for 30 minutes. This produced large grains, extending through the thickness of the strip. A specimen 4.3 cm × 0.238 cm × 0.043 cm was sawed from the portion where the grains were largest. About half this strip consists of one large grain. Another grain is about half as large, and altogether there cannot be more than twenty-five grains. That such large grains really represent single crystals was shown by examining the etching pits, which were alike over any one grain, and was checked by making a Laue photograph.[11] The records taken with this specimen do not differ more from those taken with the specimen first described than is to be expected in view of its shortness and consequent smaller change of induction for a given increment of magnetizing field. The impulses occur in the same random manner and are of about the same general intensity and number as before. A single grain, therefore, can hardly be conceived of as changing its magnetization as a unit.

If a dipole be instantaneously created in the interior of a conducting mass, the sudden growth of the magnetic field will set up eddy currents. The field of these currents will oppose the growing field of the dipole. A solenoid surrounding such a system would have a voltage pulse generated in it, consisting of an instantaneous rise to a finite maximum (initial) value and a gradual decay. The time of decay would depend on the time constant of the eddy current circuit. A paper by Wwedensky[12] enables a rough approximation of the time of the decay of the Barkhausen impulses to be made. Wwedensky considers an infinite cylinder of magnetic material, whose permeability is constant, subjected to the influence of a

[11] By Dr. R. M. Bozorth of these laboratories.

[12] B. Wwedensky, Ann. der Phys. (4) **64**, 609 (1921)

longitudinal magnetic field whose intensity changes instantaneously to a higher or lower value. He obtains an expression for the reduced mean flux through the cylinder as a function of time. Defining a sort of rough time constant as the time, τ', for the growing flux to reach $(1-1/\epsilon)$th of its final value, and using Wwedensky's expression, a silicon steel wire of 0.05 cm radius (permeability 400, conductivity 0.02×10^{-3} e.m.u.) has $\tau'=3\times10^{-5}$ sec. A lower value of permeability or smaller radius would give a smaller value of τ'.

The impulses of Table II were produced in a strip only 0.043 cm thick and the corresponding value of τ' although not directly calculable may be 10^{-5} sec. or even less. An attempt was made to detect the effect of eddy currents by comparing records taken with two specimens of silicon (3%) steel using fresh damping fluid in the oscillograph. One specimen was a wire 0.16 cm in diameter; the other was a tape 0.30 cm wide and 0.048 cm thick. Under these experimental conditions it was observed that the impulses were somewhat broadened in the case of the wire. The time constant of the deflection as measured on the oscillograph did not in either case differ much from 10^{-4} sec. The energy dissipated by such transient eddy currents is to be considered inseparable from the process of magnetization and not, like that dissipated by the eddy currents ordinarily considered, a function of the frequency of alternating magnetizing fields. There should, therefore, be a difference between the energy loss in low frequency magnetization as measured by calorimetric methods and that measured by plotting stationary values of induction and magnetizing field, in addition to the difference ascribable to eddy currents of the same frequency as the magnetizing field.

Conclusion

The above results favor the theory advanced by Barkhausen and upheld by Van der Pol that magnetic materials magnetize discontinuously, but do not substantiate the belief that each discontinuity is due to a single crystal. What determines the selection, number, and grouping of atoms which may at any time form a simultaneously magnetizable unit, and what determines the order of response of such units to an externally applied field, are not apparent, nor can any conclusion be drawn as to whether a given unit magnetizes to saturation in one jump or in several. That there should be any continuous variation of magnetization between the discontinuities seems unlikely. The apparently smooth parts of the curve observed by Van der Pol with the fluxmeter can be considered as regions in which a large number of small discontinuities occur, that is, the phenomenon can be regarded as too fine-grained for the method of obser-

vation. Little justification has been found for the theory proposed by Gerlach and Lertes and supported by Zschiesche. Their adoption of a complicated and somewhat obscure explanation depends mainly on their failure to detect an effect in iron powder and divided iron bars. With iron powder of the sort here examined even saturation of individual grains would give discontinuities too small and too numerous to give effects measurable either by their method or by that here described. It seems best to adopt the simpler notion that for some reason, limited portions of the material magnetize by jumps, either partially or to saturation, and to leave open the question of what determines the size and shape of these portions in various materials and in various specimens of the same material which have undergone different treatments. The identification of these individuals and a better understanding of their behavior would give microscopic detail to pictures of the mechanism of magnetization.

RESEARCH LABORATORIES OF THE
AMERICAN TELEPHONE AND TELEGRAPH COMPANY
AND THE WESTERN ELECTRIC COMPANY, INCORPORATED.
April 21, 1924.

ACCURATE DETERMINATIONS OF THE SPEED OF SOUND IN SEA WATER*

By E. A. Eckhardt

Abstract

Sound and radio signals were transmitted simultaneously to receiving apparatus at a distance of over 60 miles. The distances were determined to within 6 ft and the times to within .001 sec. so that the mean of 8 independent values obtained within 10 minutes is accurate to about 1 in 50,000. With practically no tide, the mean result was 4896.1 ft/sec. or 1492.3 m/sec. The method is sufficiently accurate to determine the variation in the speed of sound with temperature or salinity or to measure the component of tidal flow parallel to the sound path. The practical limitations imposed on distance determinations by the use of sound signals in sea water are thus determinable.

THE values of the speed of sound in sea water to be reported here are believed to be unique as to accuracy as well as interesting from other points of view. They were obtained incidentally during a program of work having other objectives. This fact will explain the limited number of the observations, all of which were made on a single afternoon.

The method was essentially that described by Stephenson[1] except for modifications in the distance determinations which were made necessary by the much greater distances involved. A sound and a radio signal were transmitted simultaneously at a point approximately 53 nautical miles (61 miles) from the receiving apparatus. The latter served to measure the time interval between the arrival of the radio and sound signals. This time interval was of the order of 60 sec. and was measurable to the nearest .001 sec. The distance between the sound transmitter and the sound receivers was something over 321,000 ft. The end points of this distance were located on the map to within 6 ft. The resulting speed determinations should therefore be accurate to better than one part in 25,000.

The reception and recording of the signals was accomplished by the Subaqueous Sound Ranging Section of the Coast Artillery Corps at Fort H. G. Wright who also furnished the map coordinates of the hydrophone positions. The sound and radio signals were transmitted from the steamer *Guide* of the U.S. Coast and Geodetic Survey. The sound signals

* Published by permission of the Director of the U. S. Coast and Geodetic Survey, the Chief of Coast Artillery and the Director of the U. S. Bureau of Standards.

[1] E. B. Stephenson, Phys. Rev. **21**, 181 (1923)

were made by the explosion of bombs which were suspended at a depth of 42 feet vertically below the center of a pyramidal target of a type commonly used by the Sound Ranging Section in similar work. The bombs contained about 3½ pounds of T.N.T. and were fired electrically from the 110 volt supply circuit of the ship. The operation of a single switch accomplished both the firing of the bomb and the transmission of the radio signal. It had been previously demonstrated by experiment that with the arrangement used the closing of the circuit, the detonation of the bomb, and the transmission of the radio signal are simultaneous to within .001 sec.

The position of the target was obtained at the instant of firing. Observing parties with theodolites were detached from the *Guide* and occupied stations adjacent to two lighthouses the positions of which were known from previous triangulations. They sighted on the bomb target and saw it tilt when the bomb exploded. The distances "lighthouse to ship" were of the order of five miles. As the lighthouses were occupied eccentrically the observations were reduced to center. The resultant positions of target are comparable in accuracy with that of intersection stations in tertiary triangulation. The distance from each position of the target to each hydrophone was computed by geodetic methods.

The *Guide* was anchored in about 100 feet of water and the water intervening between it and the receiving hydrophones ranged from 90 to 120 feet in depth.

The data follow.

TABLE I

Measurements of speed of sound in sea water, made Nov. 1, 1923.
Temperature of water at position of boat at depth 42 ft = 13°C

Bomb	Time	Hydrophone	Distance	Time	Speed	Speed
#1	1.55p.m.	H_3	321,900 ft	65.659 sec.	4902.6 ft/sec.	1494.3m/sec.
		H_4	323,973	66.090	4902.0	1494.1
#2	3.57p.m.	H_1	294,483	60.146	4896.12	1492.34
		H_2	297,879	60.840	4896.12	1492.33
		H_3	321,741	65.714	4896.06	1492.32
		H_4	323,808	66.135	4896.18	1492.36
#3	4.06p.m.	H_1	294,498	60.148	4896.21	1492.37
		H_2	297,894	60.843	4896.12	1492.34
		H_3	321,756	65.719	4896.00	1492.30
		H_4	323,826	66.138	4896.21	1492.37

Average velocity for Bomb #1 = 4902.3 ft/sec. (1494.2 m/sec.)
Average velocity for Bombs #2 and 3 = 4896.12 " (1492.34 ")

It is quite obvious that the data for distance and time given in the above tables permit the computation of the sound speed between the

hydrophones, leaving out of consideration entirely the radio signal accompanying the detonation of the bomb. In this way the speed over approximately the last 30,000 ft of the 300,000 ft path is obtained. The results are assembled in the following table.

TABLE II

Inter-hydrophone sound speeds

Bomb	Hydrophones	Distance	Time	Speed	Speed
#2	H_1-H_3	27,258 ft	5.568 sec	4895.4 ft/sec.	1492.1 m/sec.
	H_1-H_4	29,328	5.989	4893.9	1491.7
	H_2-H_3	23,862	4.874	4895.4	1492.1
	H_2-H_4	25,929	5.295	4897.2	1492.7
#3	H_1-H_3	27,258	5.571	4893.0	1491.4
	H_1-H_4	29,328	5.990	4896.0	1492.3
	H_2-H_3	23,862	4.876	4893.6	1491.6
	H_2-H_4	25,932	5.295	4897.2	1492.7

Average speed from inter-hydrophone measurements 4895.1 ft/sec. or 1492.1 m/sec.

While the speeds obtained from bombs No. 2 and No. 3 are in satisfactory agreement, that obtained from bomb No. 1 differs from them by approximately 6 ft/sec. This difference can be due only to a real change and not to experimental errors. By referring to a tide table it was found that at the time bomb No. 1 was fired there was a fairly strong tidal current running in the direction of sound propagation, while several hours later when bombs No. 2 and No. 3 were fired the tide was practically at a standstill. An interval of only nine minutes separated the firing of bombs No. 2 and 3. Tidal conditions for them were therefore substantially the same. The speed obtained for bomb No. 1 might therefore be taken as the speed of sound with respect to the medium plus the speed of flow of the water in the direction of propagation of the sound. The excess of 6 ft/sec. would be accounted for by a tidal current of 4.1 miles per hour, if it is assumed that the speeds are simply additive.

The Division of Tides and Currents of the U. S. Coast and Geodetic Survey has determined from data at hand what the current speeds at the time might have been and finds for the time of bomb No. 1 an average current of 1.8 miles per hour, for bomb No. 2 0.6 miles per hour and for bomb No. 3 0.36 miles per hour, all in the direction of sound propagation. While these figures are deduced from an average of actual observations, the deviations from the mean are all too small to make a current of 4.1 miles per hour probable. A large part of the 6 ft per second discrepancy between the two speed values is undoubtedly attributable to the water currents, but further work to clear up the whole discrepancy seems desirable.

The revived interest in the speed of sound is largely due to recent activity in applying it to the measurement of distances. When sea water is the medium involved, the currents, as well as inhomogeneities due to temperature and other causes limit the accuracy attainable. If speeds which are independent of the effects of flow of the water are wanted this may be accomplished by having duplicate equipment at each station and observing the time of travel for the sound in each direction. By sending the signals in both directions in rapid succession the effects of motion of the water may be completely eliminated.

DIVISION OF HYDROGRAPHY,
U. S. COAST AND GEODETIC SURVEY

SUBAQUEOUS SOUND RANGING SECT
COAST ARTILLERY CORPS,
FORT H. G. WRIGHT

SOUND LABORATORY
BUREAU OF STANDARDS

June 25, 1924

ON THE PARTIALS OF A PIANOFORTE STRING STRUCK BY AN ELASTIC HAMMER

By R. N. Ghosh

Abstract

Dynamics of a stretched string struck by an elastic hammer.—Neglecting small terms in the partial differential equation of motion, a solution is obtained involving the mass, elasticity and impact speed of the hammer and the distance of impact a from the nearer fixed end. Also approximate expressions are obtained for the pressure exerted on the string and for the duration of impact. The final solution for the displacement contains besides the Helmholtz term, three others. For a given a, the Helmholtz term disappears for partials with periods 2, 2/3, 2/5, etc., but not the other terms; this explains the observed existence of these partials with feeble intensity. The equation shows how the relative strengths of the partials depend on the various factors. If the duration of impact is half the period of vibration of a component, that component will become very prominent. For a given velocity of impact, the resulting amplitude increases with a.

Introduction

HELMHOLTZ built up his classical theory[1] supposing that the duration of impact was small compared to the period of vibration of the string, that it was not displaced appreciably from its equilibrium position, and that the hammer was elastic. Consequently he found that the duration of impact was independent of the striking point. Kaufmann in the year 1895 showed however that the duration of impact was not independent of the striking point, and gave an approximate theory for the calculation of the duration of impact when the striking distance was small, on the supposition (1) that the string was appreciably displaced from its equilibrium position, and (2) that the hammer was perfectly hard. The theory has since been extended by Raman and Banerji,[3] and also by P. Das[4] for the calculation of the duration of impact at any point of the string.

Recently it has been shown[5] that the observed values of the duration of impact do not agree with those calculated from Kaufmann's theory, at least in the middle octaves of the piano where the hammer is elastic. In the higher octaves where the hammer is almost hard Kaufmann's theory is partly true. Further it has been shown that his theory can be easily extended to the case of an elastic hammer.

[1] Helmholtz, Sensation of Tone (Ellis trans.) page 380
[2] Kaufmann, Ann. der Physik, **54,** 675 (1895)
[3] Raman and Banerji, Proc. Roy Soc. **97,** 99
[4] P. Das, Proc. Indian Asso. Calcutta **7,** 13
[5] Ghosh, Indian Science Congress Jan. 1922

In the present paper an attempt is made to point out the differences in the relative strength of the partials in the cases stated above. The main results are, (1) that the partials whose periods are 2/3, 2/5, 2/7, etc. of the duration of impact do not completely disappear; (2) that the strengths depend upon the striking point; and (3) that certain undesirable partials may become prominent under some circumstances—depending upon the striking point, the elasticity and the mass of the hammer.

Law of Motion of the String

Let: l = length of the string;
a = the striking distance from the nearer fixed end;
M = the effective mass of hammer;
μ = the strength of the elastic spring of the hammer felt;
$c = \sqrt{T/\rho}$ = velocity of transverse wave in the string;
y = the displacement of the string;
y_0 = the displacement of the center of mass of the hammer;
ξ = the compression of the hammer felt.

The equation of motion of the hammer is given by

$$Md^2y_0/dt^2 = -p \tag{1}$$

where p represents the pressure upon the hammer.

$$y_0 = y+\xi\,, \quad \text{and} \quad p=\mu\xi \tag{2}$$

and
$$T\left\{\frac{y}{a}+\frac{1}{c}\frac{dy}{dt}\right\}-p = p\frac{d^2y}{dt^2}\delta s \text{ (approx.)}. \tag{3}$$

Eq. (3) is derived on the assumption that the small part a of the string included between the striking point and the fixed end, follows the motion of the hammer, i.e. the wave takes very little time to reach the end from the striking point. Eliminating p from (1) (2) and (3) we obtain

$$-\frac{M\rho\delta s}{\mu}\frac{d^2y}{dt^2}+\frac{T}{\mu c}\frac{d^3y}{dt^3}+\frac{d^4y}{dt^4}\left\{M+\frac{MT}{\mu a}-\rho\delta s\right\} + \frac{T}{c}\frac{dy}{dt}+\frac{Ty}{a}=0\,. \tag{4}$$

A complete solution of (4) is difficult and even if it were obtained, it would be too complicated to be of any practical value; hence we shall neglect the first two terms which are very small compared to the others; then (4) reduces to

$$M\left(1+\frac{T}{\mu a}\right)\frac{d^2y}{dt^2}+\frac{T}{c}\frac{dy}{dt}+\frac{Ty}{a}=0\,. \tag{5}$$

Since T/c is small, the solution of (5) is obtained in the form

$$y = Ae^{-\frac{1}{2}kt}\sin qt \tag{6}$$

$$k = \frac{T}{Mc}\left(\frac{\mu a}{T+\mu a}\right) \tag{7}$$

$$q^2 = \frac{T}{aM}\left(\frac{\mu a}{T+\mu a}\right) - \tfrac{1}{4}k^2 . \tag{8}$$

If v_0 is the velocity of impact of the hammer at $t=0$, then A is given by

$$A = \frac{v_0}{q}\left\{1+\frac{T}{\mu}\left(\frac{1}{a}-\frac{k}{2c}\right)\right\}^{-1} . \tag{9}$$

When the hammer is perfectly hard, μ becomes infinite; then k becomes independent of the striking distance and q^2 comes out

$$q_1^2 = T/aM - \tfrac{1}{4}(T/Mc)^2. \tag{10}$$

The duration of impact is approximatly equal to π/q. It is observed from (8) and (10) that the duration of impact in the case of an elastic hammer is greater than with a hard hammer. In both cases however it increases with increasing distance of the striking point from an end.

The amplitude of the resulting motion of the string is given by Eq. (9). It is proportional to the impinging velocity of the hammer and increases as a increases.

The Force on the String

From (3) and (6) p is easily found to be

$$p = ATe^{-\frac{1}{2}kt}[\sin qt/a + (R/c)\cos(qt+\phi)] \tag{11}$$

where $R=\sqrt{q^2+k^2/4}$ and $\tan\phi = k/2q$.

From (11) it is observed that the pressure on the string is not zero at the beginning of the impact but is equal to TAq/c, and at the end of the interval π/q its magnitude is again the same; it becomes zero after a short interval δt obtained by putting $t=(\pi/q+\delta t)$ in Eq. (11). The fact that the pressure is not zero at the beginning of the impact is also confirmed by Raman's curves[6] where the hammer is considered to be perfectly hard. The interval δt however is very small; it is approximately given by a/c. Hence we shall for simplicity consider the interval π/q to be the duration of impact. This amounts to neglecting the second term in Eq. (11). Then p is given by

$$p = (AT/a)e^{-\frac{1}{2}kt}\sin qt. \tag{12}$$

[6] Raman and Banerji, loc. cit.[3], Figs. 1 and 2

Physically the neglect of the term $(R/c)\cos(qt+\phi)$ amounts to the assumption that the disturbance has traveled a long distance along the string but has not yet been reflected from the farther end, i.e. the duration of impact is smaller than the time taken by the wave to reach the farther end of the string; so that if we continually increase a, our approximations will be sufficient as long as the above condition is satisfied.

Calculation of the Partials

The equation of motion of a point on the string in terms of normal co-ordinates is given by[7]

$$\frac{d^2\phi_s}{dt^2}+\left(\frac{s\pi c}{l}\right)^2\phi_s=\frac{2}{l\rho}\Phi_s \tag{13}$$

where Φ is the generalised component of force, and

$$\Phi_s=A\sin(s\pi a/l)e^{-\frac{1}{2}kt'}\sin qt' .$$

If ϕ_s and $d\phi_s/dt$ are both zero initially then for $t>\pi/q$

$$\phi_s=\frac{2A}{nl\rho}\sin\frac{\pi sa}{l}\int_0^{\pi/q}e^{-\frac{1}{2}kt'}\sin n(t-t')\sin qt'dt$$

$$=\frac{2A}{nl\rho}\sin\frac{s\pi a}{l}\left[\frac{2q\cos(nq/2\pi)\sin n(t-\pi/2q)}{q^2-n^2}\right. \tag{14}$$

$$-\frac{\frac{1}{2}k\sin(n\pi/2q)\cos n(t+\pi/2q)}{(q-n)2}+\frac{k\pi\sin n(t+\pi/q)}{4q(q-n)}$$

$$\left.-\frac{k\pi\sin n(t-\pi/q)}{4q(q+n)}\right]\text{(approx.)} \tag{15}$$

where $n=s\pi c/l$.

The final solution for the displacement y becomes, when we substitute the value of n, to the same order of approximation,

$$y=\frac{2Al^2}{\rho\pi c}\sum_1^\infty\frac{\sin(s\pi a/l)\sin(s\pi c/l)}{s(lq-s\pi c)}\left\{\frac{2q\cos(s\pi^2c/2lq)\sin(s\pi c/l)(t-\pi/2q)}{(lq+s\pi c)}\right.$$

$$-\frac{\frac{1}{2}k\sin(s\pi^2c/2lq)\cos\{s\pi c/l(t+\pi/2q)\}}{(lq-s\pi c)}+\frac{\pi k}{4ql}\sin\frac{s\pi c}{l}(t+\pi/q)$$

$$\left.-\frac{k\pi\sin(s\pi c/l)(t-\pi/q)}{4ql(lq+s\pi c)/(lq-s\pi c)}\right\} \tag{16}$$

[7] Rayleigh, Theory of Sound, Vol 1, Art. 128

Equation (15) shows besides the Helmholtz term three other terms which appear on account of the finite displacement of the string, but on account of the presence of k their magnitudes are small. The convergence of the first two terms is of the order s^{-3}, but the last two terms converge slowly. They increase if k increases, that is when the striking distance increases. This partly explains the choice of the striking point which is about 1/7th the length of the string. Whenever the striking distance is large, the higher partials become prominent and disturb the musical quality of the note.

For a given striking distance the first term disappears for those partials whose periods are 2/1, 2/3, 2/5, 2/7 etc. of the duration of impact, but the other three terms do not, hence their existence with feeble intensity is explained. When $\sin(s\pi a/l)=0$ (i.e. the point struck is the first node of the sth component) then $a=l/s$, and the component will be absent; but if at the same time the duration of impact $\pi/q=l/cs$, then $\sin(s\pi^2c/2lq)=1$ and $\cos(s\pi^2c/2lq)=0$, and the first term and the last two terms disappear, but the second term does not and its magnitude is then approximately given by $k/2s$ which is small and decreases with the order of the component. So we find that a component having a node at the point struck is not entirely absent if the duration of impact is half the period of vibration of the component. Hipkins and others have found this to be the case experimentally.[8]

A given component will converge more rapidly the greater the value of q, which increases as a decreases (see Eq. 8), hence a must be small so that the higher undesirable partial may be very feeble. It is further observed that the magnitude of the second term is inversely proportional to $(lq-s\pi c)^2$ hence when $q=s\pi c/l$, that is when the duration of impact is equal to half the period of vibration of that component, it will become very prominent.

A is not constant throughout all the length of the string for the same velocity of impact of the hammer. It changes with q and therefore increases as a the striking distance increases.

PHYSICS DEPARTMENT,
ALLAHABAD UNIVERSITY,
ALLAHABAD, INDIA.
December 7, 1923.

[8] Helmholtz, loc. cit.[1] pp. 76, 77

BOOK REVIEWS

Color and Methods of Color Reproduction. By MARTIN AND GAMBLE.—This book contains a great deal of interesting and useful information on color. The exposition of the more elementary portions is clear, and it is evident from the frequent references to easily observable phenomena in nature dependent on color, and to works of art, that the writer uses and enjoys his eyes. The chapters on colored light and coloring materials provide useful, practical information on methods of producing color, both in the laboratory and in the arts. Detailed information on the chemical characteristics and the permanence of various coloring materials is something that the physicist working with color often wants and usually has difficulty in finding.

Chapters on color vision, color blindness and color in human experience give an interesting general review of the physiological and psychological aspects of color, treated in a fairly popular manner. Photometry and color measurement are considered at some length, but not entirely satisfactorily, since a number of erroneous statements, particularly of experimental findings in the field of intermittent vision, indicate that the author's acquaintance with this part of the subject has been gathered from rather hasty and undigested reading. The method given of calculating hue, luminosity and purity from spectrophotometric data does not show acquaintance with the more recent work on color measurement transformations. In particular the "saturation" arrived at is not the saturation with respect to the spectrum, but to a non-existent "two sensation" spectrum.

Two chapters of a very popular and non-technical character by the second author, on "Color printing" and "Color photography" would have been of greater assistance to a clear presentation of three-color printing processes if placed in inverted order. The scientific basis of these processes and of three-color photography are not clearly grasped by the authors. "Red, yellow and blue" are given as the primary printing colors, and the author of the first part of the book, while in one place correctly stating that the ideal subtractive primaries would be "peacock blue," "chrome-yellow" and "cerise," does not stand by this consistently. His colored plates illustrative of additive and subtractive color mixture (p. 17) show in each case the complementary of red as blue. The foot-note to page 166 which endeavors to make clear the three color additive process of color photography by optical superposition of the three color images is unfortunate in the parenthesis "both for taking and viewing." The simple principle that the transmission of the "taking" screens should copy the color mixture curves of the spectrum while the viewing screens should be the monochromatic primaries used for determining the color mixture curves was clearly enunciated at least thirty years ago.

These illustrations are sufficient to show that, interesting and valuable as this book is in many ways, it nevertheless cannot satisfy the student who wants sound detailed information. It is only fair to state however that a comparatively small number of text changes would suffice to correct the errors which have been noted. Such changes would make a second edition a much more valuable book.—Pp. 187, 73 fig., Van Nostrand, 1923. HERBERT E. IVES

The Elementary Principles of Lighting and Photometry. By J. W. T. WALSH.—To quote the author, this book "is mainly intended to give a description of the nature and amount of the illumination required for different purposes, and of the way in which the desired result may be attained and its attainment checked by photometric measure-

ment." Thus while a large part of the subject matter is divided into "Indoor Lighting," "Industrial and School Lighting," "Outdoor Lighting," "Light Projection" and such practical illumination matters, these are well introduced by several chapters dealing with "Light, Vision and the Eye," and an outline of photometry.

The chapters dealing with photometry present a well balanced review of the elementary aspects of the subject. Good judgment is shown in the relative space given to subjects of academic rather than practical interest, such for instance as that of fundamental photometric standards. The treatment of color and color difference photometry is not so satisfying. The rating of illuminants by average candle power rather than lumens follows an English practice which is some years behind our own. Omission of any discussion of luminous efficiencies or any adequate treatment of the physics of light production is to be regretted. A carefully chosen bibliography at the end, together with appendices giving the most authoritative definitions of photometric quantities will assist the reader to fill in these gaps.

The guidance offered to the practical problems of lighting is good. "The important factor in seeing is the brightness of the thing looked at and this is the product of the illumination and the reflection ratio," but the qualifications to this introduced by the effects of bright lights in the field of view, and glare in general, are well covered. Specific treatments of lighting schemes for individual rooms in the house should prove of real help to people of the class to whom this book is addressed, namely those possessed of an elementary scientific and technical education, who are, let us say, not lighting engineers but nevertheless are desirous of lighting their living rooms safely and sanely.—

HERBERT E. IVES

Light and Color in Advertising and Merchandising. By M. LUCKEISH.—This book does not purport to be a scientific treatise on light and color. To quote the author "it seems unwise to go into these tedious details here." The chief interest of the book to the physicist will be the revelation it gives of how far reaching may be the reactions, on the human complex, of certain rather simple physical variables. Thus, starting with radiant energy which can be varied physically merely in intensity, frequency and direction of propagation, our visual apparatus attributes it to this or that localized material object, and adds the quality of color. Our mental experiences differentiate between pleasant and unpleasant arrangements of the visual field of view, and sets of color associations arise, harmonious and unharmonious. These become in time part of our language and literature. Ultimately a study of how these factors may be intelligently applied by the advertising specialist, the interior decorator, and the window trimmer, comes to be written, showing how these practical men who know nothing of lumens or Angstroms, can make real money by proper use of light and color.

The book consists essentially of a popular review of what is known of the psychology of light and color, with numerous specific commercial applications. A large number of colored illustrations make the book a very attractive one, even to the reader whose interest in the subject matter may be slight.—Pp. xvi+268. 38 illustrations, Van Nostrand, 1923.

HERBERT E. IVES